



Analysis of the non classical class I genes of the MHC in swine

Hu Rui

► To cite this version:

Hu Rui. Analysis of the non classical class I genes of the MHC in swine. Life Sciences [q-bio]. Université de Versailles-Saint Quentin en Yvelines, 2011. English. NNT: . tel-00609548

HAL Id: tel-00609548

<https://theses.hal.science/tel-00609548>

Submitted on 19 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN-YVELINES

Ecole Doctorale des Génomes Aux Organismes (DGAO)

UMR de Génétique Animale et Biologie Intégrative
INRA, Département de Génétique Animale
AgroParisTech, Département Sciences de la Vie et Santé

THESE DE DOCTORAT
DE L'UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN YVELINES
Spécialité : Génétique moléculaire

Présentée par : Rui HU

Pour obtenir le grade de Docteur de l'Université de Versailles Saint-Quentin-en-Yvelines

**Etude des gènes de classe I non classiques du complexe majeur
d'histocompatibilité chez le porc**

Analysis of the non classical class I genes of the MHC in swine

Soutenue le : vendredi 13 mai 2011

Directeur de thèse: Dr Claire Rogel-Gaillard, INRA

Devant le jury composé de :

Président : Pr Bernard Mignotte, Université de Versailles Saint Quentin en Yvelines

Rapporteurs: Dr Keith BALLINGALL, Moredun Research Institute, Scotland, UK
Dr Philippe Le BOUTEILLER, INSERM, Toulouse, France

Examineurs : Dr Isabelle SCHWARTZ-CORNIL, INRA, Jouy-en-Josas, France
Dr Isabelle HUE, INRA, Jouy-en-Josas, France
Dr Patrick CHARDON, INRA, Jouy-en-Josas, France

TABLE OF CONTENTS

Résumé	1
Avant-propos.....	5
Introduction.....	9
1- History of MHC research	9
1.1- In mouse and human	9
1.2- In swine.....	10
2- The MHC locus in mammals.....	11
2.1- Chromosomal mapping.....	11
2.2- Overall organization and size.....	11
2.3- The MHC class III region: a well conserved segment	13
2.4- The MHC class II region	14
2.4.1- HLA class II region.....	14
2.4.2- SLA class II region.....	14
2.4.3- BoLA and OLA class II	15
2.5- The MHC class I region.....	15
2.5.1- HLA class I region	15
2.5.2- SLA class I region	16
3- MHC class Ia genes and molecules.....	17
3.1- Organization and gene structure.....	17
3.2- Expression.....	19
3.3- Functions	20
3.3.1- Peptide presentation to cytotoxic CD8+ T lymphocytes	20
3.3.2- Modulation of Natural Killer (NK) cells	22
3.4- Polymorphism	23

4- MHC class Ib genes and molecules.....	26
4.1- Gene structure	26
4.2- HLA class Ib	27
4.2.1- HLA-E	27
4.2.2- HLA-F	29
4.2.3- HLA-G	32
4.3- SLA class Ib	36
4.4- Polymorphism	37
5- Aim of the work	40
Materials and Methods	43
1- Transcription studies.....	43
1.1- Animals and samples	43
1.2- RNA extraction	43
1.3- Primer design.....	44
1.4- Reverse Transcription.....	46
1.5- PCR amplification.....	46
1.6- DNA fragment purification.....	47
1.6.1- Purification of DNA from agarose gels	47
1.6.2- Direct purification from PCR	47
1.7- PCR fragment cloning and sequencing.....	48
1.7.1- Ligation reaction	48
1.7.2- Preparation of electro-competent bacteria.....	48
1.7.3- Transformation of competent bacteria	49
1.7.4- Screening of recombinant bacteria	50
1.7.5- Plasmid DNA preparation	50
1.7.6- DNA sequencing	50

1.8- Sequence analysis	51
2- Protein studies.....	52
2.1- Protein structure prediction.....	52
2.2- Construction of expression vectors	52
2.2.2- Vectors for transfection in mammalian cells.....	53
2.2.3- Vectors for transfection in drosophila cells	53
2.3- Large scale preparation of plasmid DNA.....	55
2.4- Cell lines and transfection	56
2.4.1- Insect cells: the drosophila Schneider 2 cells (SC2)	56
2.4.1.a- Cell culture conditions	56
2.4.1.b- Transient transfection of SC2 cells.....	57
2.4.1.c- Stable transfection of SC2 cells.....	58
2.4.2- Mammalian cells: porcine PK15 cells	58
2.5- RNA extraction from transfected cells	59
2.6- Monoclonal antibody production specific for the SLA-Ib molecules	59
2.7- Western blot analysis for detection of expressed proteins in SC2 cells.....	59
2.8- Fluorescent Activating Cell Sorting (FACS) analysis	60
2.8.1- Detection of molecules on the cell surface	60
2.8.2- Detection of molecules in the cytoplasm by cell permeabilization	61
3- Polymorphism studies.....	62
3.1- Primer sets for long range PCRs specific for the SLA-Ib genes	62
3.2- Animals.....	63
3.3- Long Range PCR and cloning.....	63
3.4- DNA preparation and sequencing	64
3.5- Sequence analysis	64
3.6- Pyrosequencing.....	66

Results	69
1- SLA Ib transcription	71
1.1- Characterization of SLA-Ib transcripts.....	71
1.1.1- Identification of alternative transcription for SLA-6 and -7 but not SLA-8	71
1.1.2- SLA-6: five transcripts	72
1.1.2.b- SLA-6-1 to SLA-6-4 transcripts: alternative spliced variants	73
1.1.3- SLA-7: seven transcripts	74
1.1.3.a- SLA-7-001 and SLA-7: two different transcripts that stand for the full coding sequence	76
1.1.3.b- SLA-7-1 to SLA-7-5: alternative RNA variants with a complex pattern.....	76
1.1.4- SLA-8: a unique full length transcript.....	78
1.1.5- Validation of two splice sites by genomic DNA amplification of SLA Ib.....	78
1.2- Tissue specificity	80
1.2.1- Relative expression of SLA-Ib genes and comparison with SLA-Ia genes	80
1.2.2- Expression of the transcript variants in various tissues	81
1.2.2.a- SLA-6	81
1.2.2.b- SLA-7	83
2- SLA Ib protein.....	87
2.1- Protein structure	87
2.1.1- Predicted SLA-6 protein isoforms	87
2.1.2- Predicted SLA-7 protein isoforms	89
2.1.3- Predicted SLA-8 protein	91
2.1.4- Conclusion on protein prediction	91
2.2- Protein expression.....	92
2.2.1- Experimental design of the experiments	92
2.2.2- Construction of the expression vectors	93
2.2.2.a- Expression vectors for expression in mammalian cells	93
2.2.2.b- Expression vectors for expression in Drosophila cells.....	94

2.2.3- Protein expression in drosophila cells	95
2.2.3.a- Detection of target RNAs in transfected cells	95
2.2.3.b- Detection of target proteins in transfected cells	96
2.2.4- Protein expression in PK15 cells and antibody screening	99
3- SLA Ib polymorphism.....	105
3.1- cDNA sequencing: identification of coding SNPs (cSNPs)	105
3.1.1- SLA-6	105
3.1.2- SLA-7	106
3.1.3- SLA-8	107
3.2- Characterization of the nucleotide polymorphism at the whole gene level	107
3.2.1- Experimental design	107
3.2.2- SLA-7 polymorphism	110
3.3- Copy Number Variation (CNV) of SLA Ib genes	115
3.3.1- SLA-7 SNP haplotypes.....	116
3.3.2- SNP quantification by pyrosequencing	116
Discussion.....	125
1- Polymorphism of SLA-Ib genes: characterization of SNPs and CNVs	125
2- A splicing pattern with various complexities according to genes.....	129
3- Tissue specificity of SLA-Ib gene transcription and variations observed between MeLiM and Large White animals	132
4- Comparison of SLA Ib to HLA Ib genes	135
5- Properties of the putative proteins encoded by SLA-Ib genes	136
Conclusion	139
Reference	140
Publication # 1	151

Publication # 2	171
Abstracts and posters.....	184
Supplementary files	189
Abstract / Résumé.....	211

LIST OF FIGURES

Figure I: General organization of MHC in human, pig and ovine	12
Figure II: The genomic organization of the pig and human MHC class I regions.....	17
Figure III: The structure of MHC class I molecules	18
Figure IV: MHC class Ia gene structure	19
Figure V: MHC class I antigen presentation procedure	22
Figure VI: Comparative gene structure of SLA-Ia and Ib genes	26
Figure VII: HLA-E transcripts and protein structure	27
Figure VIII: HLA-F transcripts (A) and predicted protein structures (B).....	30
Figure IX: Alternatively HLA-G transcripts and protein isoform structure	34
Figure X: HLA-G immune cell interaction	35
Figure XI: Restriction map of the plasmid pVAX1-SLA6	53
Figure XII: Map and features of the expression vector pAc5.1/V5-His	54

Figure XIII: Cloning of cDNAs into the expression vector (pAc5.1/V5-His vector).....	55
Figure XIV: General strategy to study the polymorphism of SLA-Ib genes	62
Figure XV: PCR result of full-length cDNA of SLA Ib	72
Figure XVI: Five SLA-6 transcripts obtained from MeLiM pig	73
Figure XVII: SLA-7 transcripts obtained from MeLiM pig	75
Figure XVIII: SLA-8 transcript obtained from MeLiM pig	78
Figure XIX: Validation of splicing events by PCR amplification of SLA-6 (A) and SLA-7 (B) partial segments from genomic DNA	79
Figure XX: Detection of SLA-6 RNA variants in various tissues from MeLiM and Large White pigs by RT-PCR	82
Figure XXI: Detection of SLA-7 RNA variants in various tissues from MeLiM and Large White pigs by RT-PCR	84
Figure XXII: Detection of SLA-8 RNAs in various tissues from MeLiM and Large White pigs by RT-PCR	85
Figure XXIII: Prediction of SLA-6 protein encoded by the various splice variants.....	89
Figure XXIV: Prediction of SLA-7 protein isoforms encoded by the splice variants SLA-7 (A, B) and SLA-7-1 (C, D)	90
Figure XXV: Prediction of SLA-8 protein.....	91
Figure XXVI: The expression vectors of SLA-6, -7, -8, -1, and B2M	95

Figure XXVII: Detection of B2M transcripts by RT-PCR	96
Figure XXVIII: Cell surface expression of B2M by FACS.....	98
Figure XXIX: Detection of SLA molecules by FACS in the cytoplasm (A) or on surface (B) of PK15 cells transiently transfected with the vector pVAX1-SLA6	100
Figure XXX: Detection of SLA molecules by FACS on surface of PK15 cells transiently transfected with the vector pVAX1-SLA8.....	101
Figure XXXI: The primer positions used in primer walking sequencing.....	108
Figure XXXII: Gel electrophoresis showing amplification of SLA-6, -7 and -8 genes by long range PCRs.....	109
Figure XXXIII: Different haplotypes of SLA-7 identified in the MeLiM pig 484	116
Figure XXXIV: Primer design for pyrosequencing	119
Figure XXXV: Pyrosequencing result for SNP1535 and SNP1540 from MeLiM pig.....	120

LIST OF TABLES

Table I: SLA Ia polymorphism and haplotypes	25
Table II: Expression patterns of HLA-Ib genes inferred from EST sources.....	31
Table III: The comparison of polymorphism between class Ia and class Ib genes	38
Table IV: SLA-6 polymorphism and SLA haplotypes	39

Table V: Gene specific primers used for transcription studies	45
Table VI: Universal and internal primers used for SLA Ib transcript sequencing.....	51
Table VII: cDNA sequence features	52
Table VIII: Antibodies used for FACS analysis	61
Table IX: Primer sets designed to amplify the SLA Ib genomic sequences	63
Table X: Primers for SLA Ib gene sequencing	64
Table XI: Primers for pyrosequencing PCR and sequencing.....	66
Table XII: The mixture for PCR amplification.....	68
Table XIII: Characteristics of the inserts sub-cloned into expression vectors.....	94
Table XIV: Screening of anti-SLA-6 monoclonal antibodies by FACS in PK15 cells transiently transfected with the pVAX1-SLA6 vector.....	102
Table XV: Screening of anti-SLA-8 monoclonal antibodies by FACS in PK15 cells transiently transfected with the pVAX1-SLA8 vector.....	103
Table XVI: SLA-6 Coding SNP Position	106
Table XVII: SLA-7 Coding SNP position	107
Table XVIII: Number of clones selected per animal for sequencing SLA-7 and SLA-8 genes Clone information	110
Table XIX: SNP positions on SLA-7 gene	111

Table: XX: SNP position on SLA-8 gene	114
Tableau XXI: Allele frequency of SNPs at positions 3601, 1535 and 1540 by pyrosequencing	121
Tableau XXII: The relative frequency of the paralogous sequences of SNP 1535	122
Tableau XXIII: The relative frequency of the paralogous sequences of SNP1540	123

Résumé

Le complexe majeur d'histocompatibilité (CMH) a été intensément étudié depuis qu'il a été identifié comme un acteur majeur de la réponse immunitaire en distinguant les molécules du soi de celles du non soi. Le CMH est une région conservée chez les vertébrés supérieurs à mâchoires et son apparition est souvent reliée à l'émergence de la réponse immunitaire adaptative. Le CMH est sub-divisé en trois sous-régions, appelées régions de class I, II et III. Les régions de classe I et II contiennent les gènes qui codent pour les molécules d'histocompatibilité de classe I et II, respectivement. Les molécules de classe I présentent des peptides endogènes et viraux aux cellules T cytotoxiques et modulent l'activité des cellules natural killer (NK). Les molécules de classe II présentent des peptides exogènes aux cellules T helper CD4+. Outre ces gènes d'histocompatibilité, le CMH contient de nombreux autres gènes reliés ou non à la réponse immunitaire, dont les fonctions ne sont pas toutes élucidées.

Suite aux travaux de cartographie et de séquençage du CMH, il a été montré que les gènes de classe I se subdivisent en deux groupes qui distinguent les gènes de classe I classiques (Ia) qui codent pour les molécules dont les fonctions ont été décrites ci-dessus des gènes de classe I non classiques (Ib). D'après ce qui est connu pour l'espèce humaine, les molécules Ib du CMH connues sous le nom de HLA-E, -F et -G, ont des fonctions immuno-modulatrices lors de l'induction de la tolérance de la mère pour le fœtus au cours de la grossesse, notamment en modulant l'activité des cellules NK. Des fonctions de tolérance immunitaire ont également été décrites dans le cas de tumeurs solides comme des mélanomes chez l'homme. Les gènes Ia du CMH sont caractérisés par une expression quasi ubiquitaire et un niveau élevé de polymorphisme, essentiellement concentré dans les exons 2

et 3 correspondant aux domaines alpha 1 et 2 qui forment le sillon dans lequel vient se positionner le peptide présenté aux cellules T cytotoxiques. A l'inverse, les gènes Ib du CMH ont un niveau faible de polymorphisme, présentent des variations de structure dans la région qui code pour la queue intra-cytoplasmique, ont une expression tissulaire restreinte exprimés et expriment des transcrits alternatifs. L'expression et le rôle de ces gènes de classe Ib est très peu étudié chez les animaux non modèles comme le porc, alors qu'ils revêtent un intérêt majeur s'ils sont capables d'induire ou de lever une immuno-tolérance.

L'objectif de la thèse a été de caractériser l'expression et le polymorphisme des gènes SLA-6, SLA-7 et de SLA-8 qui sont les trois gènes du CMH caractérisés comme de classe I non classique chez le porc. Nous avons mené ce travail en utilisant des échantillons biologiques provenant de porcs Large White et de porcs MeLiM qui développent spontanément des mélanomes. Nos résultats ont permis de démontrer que les trois gènes sont transcrits dans une large gamme de tissus, avec une prédominance dans les tissus lymphoïdes, le système digestif et les poumons. Les transcrits SLA-6 et -7 présentent des épissages alternatifs alors qu'un unique transcrit a été identifié pour le gène SLA-8. Les trois gènes expriment des transcrits qui codent pour des protéines complètes dont la prédiction de conformation est compatible avec la présentation de peptide à la surface des cellules. SLA-6 et -7 expriment des transcrits alternatifs qui codent pour des protéines putatives de propriétés variables dont certaines pourraient être solubles, de par l'absence de domaine trans-membranaire. Le gène SLA-7 s'est révélé le plus complexe. Les études de transcription ont permis la ré-annotation du gène avec huit exons au lieu de sept comme initialement décrit. De plus, nous avons montré l'existence d'un épissage alternatif dans l'exon 4 avec l'utilisation d'un site rare d'épissage (GA-AG au lieu de GT-AG), ainsi que l'existence d'un épissage alternatif dans la

région 3' non codante, en aval du codon de terminaison. Les deux formes de transcrits dans la région 3' non codante ont été trouvées co-exprimées dans les tissus de porcs MeLiM alors que seul le transcrit non épissé a été trouvé dans les tissus de porcs Large White. Nos résultats confirment un niveau faible de polymorphisme nucléotidique et démontrent l'existence d'un nombre variable de copies du gène SLA-7 selon les races de porc: au moins deux copies du gène chez les porcs MeLiM alors qu'une copie est présente chez le porc Large White inclus dans notre étude. Ces résultats indiquent que le polymorphisme du gène SLA-7 inclue à la fois des variations de nucléotides et de nombre de copies. Afin de progresser dans l'étude de l'expression des molécules SLA-Ib, nous avons mis au point des systèmes d'expression des molécules dans des cellules épithéliales de porc et dans des cellules de drosophile et avons fait produire par une société commerciale, au cours de la thèse, des anticorps dirigés contre chacune des molécules. Les anticorps monoclonaux anti-SLA-6 et -8 sont en cours de test et reconnaissent des molécules exprimées en surface mais nous ne pouvons encore conclure quant à leur réelle spécificité. Pour conclure, nous avons produits des données nouvelles sur les gènes SLA-Ib à l'échelle de la structure génomique et de la transcription et les expériences en cours sur l'expression des protéines devraient nous permettre prochainement de conclure quant aux propriétés de ces molécules de s'exprimer à la surface des cellules. L'ensemble de ces études est une étape vers d'autres travaux qui viseront à étudier l'expression des gènes SLA-Ib au cours du développement embryonnaire, à l'interface foeto-maternelle, ainsi que dans les tumeurs en progression et régression dans le modèle biomédical MeLiM. Savoir si les molécules SLA-Ib sont les homologues fonctionnelles des molécules HLA-Ib connues pour leur fonction immuno-modulatrices est un essentiel pour comprendre les différents types de réponses immunitaires chez le porc.

Avant-propos

The Major Histocompatibility Complex (MHC) has been intensively studied since it was first identified as a major actor of individual immune response against pathogens by distinguishing self from non self. The MHC can be considered one of the most well known genomic regions among living vertebrate organisms. However, knowledge about MHC is far from completed. Sequencing data and molecular immunology as well as a growing number of genetic association studies constantly provide new insights into MHC studies and functions. The MHC was first described as a genetic system due to a high polymorphism of MHC molecules. Before genome mapping and sequencing techniques could be applied to MHC analysis, studies on MHC were carried out with cellular, serological, and immunochemical analysis methods. Those methods revealed the biochemical and immunochemical features of MHC, but the genomic information remained limited. Genetic and physical maps were further drawn and during the last decade, the locus has been fully sequenced in many species. The first complete MHC was sequenced in the human in 1999 (MHC-Sequencing-Consortium 1999) followed by the mouse in 2002 (Waterston et al. 2002), the chimpanzee MHC class I region in 2003 (Anzai et al. 2003), the rat in 2004 (Hurt et al. 2004), the pig in 2006 (Renard et al. 2006), and the cattle in 2010 (Gao et al. 2010). These new datasets contributed to the identification of numerous genes, many of which do not encode MHC molecules or have an immunity-related function.

Sequencing data have clarified the number and genomic organization of genes in each MHC gene family and have provided a refined classification of MHC class I genes into classical (Ia) and non-classical (Ib) subgroups according to genomic structure and polymorphism. A simple assessment could be that classical MHC molecules have roles attributed to the MHC

function and non-classical molecules have complementary functions not yet fully understood. Moreover, it is well established that classical genes have a universal function among species whereas non-classical genes may have either functions shared between species or species specific functions. Therefore, analyses of non-classical genes must be addressed species by species.

The pig is an important agricultural animal as well as a model animal for biomedical research (e.g. xenotransplantation). Swine is known as a highly relevant species for human disease and physiology studies due to its high degree of similarity to humans. Since the initial studies, a large research community working on the pig MHC was organized and the pig MHC research has followed the same main steps as in other species: description of a genetic system, characterization of haplotypes, physical mapping, sequencing and gene annotation. As in other species, these findings have led to the refinement of the MHC gene series and identification of non-classical genes whose functions remain to be elucidated.

Very limited data are available on the three non-classical MHC class I genes referred to as SLA-6, SLA-7 and SLA-8 in swine. It is a major issue to know whether these genes have immunity-related functions that would be pig-specific or if they are functional homologues of class Ib genes in other species. In various species, the class Ib genes were reported to differ from class Ia genes by several features that include a limited polymorphism, a modified structure of the cytoplasmic tail, tissue specificity and expression of alternative splice transcripts. Moreover, all encoded molecules do not have the same ability to present peptides on cell surface. In order to increase knowledge on MHC class Ib genes and molecules in swine, our purpose was to analyse the SLA-6, -7 and -8 genes at the genomic, RNA and protein levels. The corresponding questions are summarized as follows:

- At the genomic DNA level: are the SLA-Ib genes oligomorphic, as expected ?
- At the transcription level: i)- do SLA-6, -7 and -8 express alternative variants and if yes, which ones? ii) - in which tissues are the RNAs expressed?
- At the protein level: are the SLA-Ib molecules expressed on cell surface and if yes, do they present peptides?

In the introduction, history about the MHC discovery, MHC organization and knowledge about classical and non-classical class I genes will be summarized, mostly by comparing human and pig MHCs. The Material and Methods and Result sections have been divided in three parts related to transcription studies, protein expression, and gene polymorphism. Two articles have been accepted and are attached at the end of the manuscript. The last section aims at discussing new insights on MHC non-classical gene studies provided by the results presented in this report.

Introduction

1- History of MHC research

The MHC is a genomic region that contains a cluster of genes involved in the immune response (Horton et al. 2004). The MHC region was found in all jawed vertebrates examined to date (Kelley et al. 2005). The function of MHC molecules in presenting peptides to various T cell subsets has been identified as a crucial strategy to recognize peptides derived from foreign pathogens and to protect organisms from disease. Occurrence of MHC has often been related to the emergence of adaptive immunity as reviewed by Danchin et al (Danchin et al. 2004).

1.1- In mouse and human

MHC studies started more than 70 years ago. In 1936, Peter A. Gorer first reported the mouse antigen II as a major histocompatibility antigen that was named histocompatibility-2 (H-2) (Schutze et al. 1936). Later, studies showed that the H-2 contained a series of variable molecules and H-2 was further described as the H-2 complex. The mouse H-2 complex was the first major histocompatibility complex discovered (Klein 1986). A system analogous to the H2-complex was described in human by Jean Dausset (Dausset 1958) and further referred to as the Human Leucocyte Antigen Complex or HLA complex (Ivanyi J. and Pinter 1967).

The phenomenon of histocompatibility was identified in the middle of the 20th century when dealing with transplantation experiments aimed to grafting skin to the airmen burned after the Second World War. In the procedure of skin graft, it was found that the acceptor rejected

foreign tissue grafts. Further studies revealed that MHC variability between the donors and acceptors lead to the rejection of the donor graft being identified as non-self by the immune system of the acceptor (Auchincloss et al. 1993, Gould and Auchincloss 1999, Haeney 1995).

The role of MHC in the immune system was unknown until the early 1970s when MHC was shown to control the immune reaction by antigen presentation (Benacerr.B and Mcdevitt 1972, Snell 1976). It was also shown that MHC molecules present peptides to T cell receptors, T cells recognize antigens presented by MHC molecules in a highly restricted manner, meaning that the interaction between T cells and antigen presenting cells originates from a syngeneic background. These findings gave rise to the concept of MHC restriction and MHC-restricted T cells (Archbold et al. 2008, Zinkernagel and Doherty 1974).

1.2- In swine

The first studies in swine started in 1940s and aimed to analyze correlations between allograft acceptance or rejection and known blood groups. In 1970, it was suggested that the blood system E could correspond to the pig MHC (Ivanyi P. 1970). The pig MHC was identified by immunizing piglets from the same litters with cutaneous grafts. The serums from immunized piglets were used to identify the various allelic forms of the class I molecules responsible for immunization. In a second step, longer graft acceptance was found for donor and acceptor animals harbouring similar alleles, confirming that the identified alleles were related to the histocompatibility molecules (Vaiman et al. 1970). The pig MHC was referred to as Swine Leucocyte Antigen complex or SLA complex.

2- The MHC locus in mammals

2.1- Chromosomal mapping

The MHC has been assigned to a chromosome for almost all species for which it has been characterized (review in Kelley et al. 2005). In human, the MHC maps to chromosome 6p21.1 (Lamm and Olaisen 1985). In mouse, the H2 complex maps to chromosome 17. In ovine, the MHC complex (Ovine Leukocyte Antigen or OLA complex) maps to chromosome 20 at position q15-q23 (Dukkipati et al. 2006). In cattle, the MHC (Bovine Leukocyte Antigen or BoLA complex) maps to chromosome 23 (Brinkmeyer-Langford C. et al. 2008).

In pig, the MHC maps to chromosome 7 (Geffrotin et al. 2004) and has been shown to be interrupted by the centromere (Rabin et al. 1985), the class I and III regions mapping to the p arm (7p11) and the class II region mapping to the q arm (7q11). This feature seems unique to the pig among all species in which MHC has been characterized so far.

2.2- Overall organization and size

The MHC locus has been divided into three regions according to the gene content referred to as class I, II and III, (Figure I). The class I and II regions contain MHC class I and II genes, respectively, and the class III region is located at the junction between class I and II and has been named class III in order to split the MHC into regions with similar names. The class III region does not contain histocompatibility genes. Recent reports resulting from sequencing projects have characterized an extended MHC on both sides of the MHC locus in human (Horton et al. 2004). Both extended segments are characterized by a high gene density and

an interesting cluster of olfactory receptor genes mapped in the vicinity of extended class I region (Horton et al. 2004).

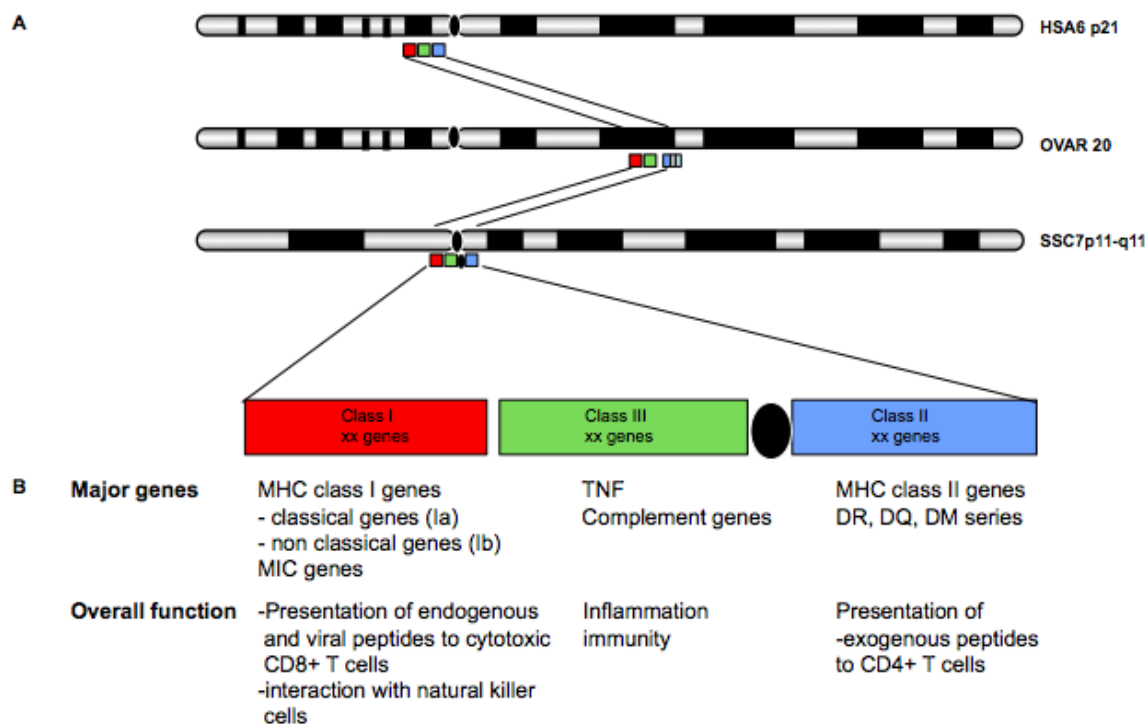


Figure I: General organization of MHC in human, pig and ovine

A. The general scheme of the MHC locus: cytogenetic position of the locus in human (HSA6p21), ovine(ovar20) and pig (SSC7p11-q11). The MHC region is divided into three sub-regions referred to as class I (red box), II (blue box), and III (green box). In human, all three classes MHC map to the long arm. In ovine, the difference from human is that the class II is divided into class IIa and class IIb. In pig, Class I and III map to the short arm while class II is divided from I and III by the centromere situating on the long arm. In the pig, B: The major genes and related functions specific to each MHC sub-region and the summary of major gene content and associated function.

The MHC locus spans several megabases (Mb) in mammalian species. In human, the HLA complex spans 3.6 Mb and 224 genes are annotated along with 128 genes predicted to be expressed (Singh-Gasson et al. 1999). By contrast, the organizational features of the MHCs

of cattle and other ruminants are unique in that class II genes occur in two segments rather than a single segment as observed in other mammalian species such as in human, mouse (Waterston et al. 2002), dog (Wagner 2003), and horse (Gustafson et al. 2003). The two segments are located about 20 cM apart and are designated class IIa and class IIb (Andersson et al. 1998, Band et al. 1998, Brinkmeyer-Langford C. L. et al. 2009, Skow et al. 1996). Class IIa is closely associated with the class I and class III regions, while class IIb is positioned closer to the centromere. The gene order of class IIb in both ovine and bovine regions show an opposite orientation relative to that of human (Gao et al. 2010). The OLA complex spans 2.4 Mb with prediction of 177 protein-coding genes (Gao et al. 2010).

The SLA complex spans 2.4 Mb with an overall organization similar to the HLA complex despite the interruption by the centromere; 151 SLA loci have been annotated and 121 predicted to be functional (Renard et al. 2006).

2.3- The MHC class III region: a well conserved segment

The HLA class III region spans 0.7 Mb in human and swine and encodes non-MHC molecules such as C2, C4 and B factor that belong to the complement system, TNF, LTA and LTB that are cytokines related with inflammation. The class III region also includes heat shock proteins (HSPs).

2.4- The MHC class II region

2.4.1- HLA class II region

The HLA class II region spans 0.9 Mb from C6orf10 to HCG24 and contains HLA-DRA, -DRB1, -DRB2, -DRB3, -DRB4, -DRB5, -DRB6, -DRB7, -DRB8, -DRB9, -DPB1, -DMA, -DMB, -DOA, and -DOB. The classical molecules include the DP, DQ and DR series, and the non-classical molecules include the DM and DO gene series. The TAP1 and 2 genes (transporter associated with antigen processing) as well as the proteasome subunits PSMB7 and PSMB9 also map to the class II sub-region.

2.4.2- SLA class II region

The SLA class II sub-region spans 417 Kilobases (Kb) from the butyrophilin like (BTNL) gene cluster close to the centromere downstream to the last annotated locus DOA (Renard et al. 2006). In total, 25 loci are annotated. A single SLA-DRA gene and five SLA-DRB loci are described. DRB1 is full length and DRB4 has a deletion in exon 1. Exons 1 and 6 are missing in DRB3, while only exon 6 is missing in DRB5 and DRB2. Four of five DRB loci are oriented and clustered in a pattern similar to that of other mammals; the remaining SLA-DRB5 locus lies on the opposite strand within the DQ–DO interval. The SLA-DQ region comprises one DQA locus and two DQB loci of which only one is functional. The number of DRB and DQB copies could vary between different SLA haplotypes, as observed in the HLA (Horton et al. 2004). The SLA DQ–DO interval also contains a putative locus (SBAB-554F3.8) and three pseudogenes (SBAB-554F3.9, SLA-DOB2, SLA-DYB) with similarities to TAP2, DO, and artiodactyl-specific DYB, respectively (Renard et al. 2006). From the most centromeric SLA-DRA gene in the class II gene cluster, the order of the expressed

SLA genes is DRB1, DQA, DQB1, DOB1, DMB, DMA and DOA. There are eight pseudogenes in the SLA class II region: SLA-DRB2, SLA-DRB3, SLA-DRB4, SLA-DRB5, SLA-DQB3, SLA-DOB2, wDYA and wDYB. The class II region also comprises TAP1 and TAP2 genes as well as PSMB7 and PSMB9 genes.

2.4.3- BoLA and OLA class II

In cattle and sheep, the class II DQ-DO interval is split into two subregions separated by 17–30 cM (Amills et al. 1998, Jarrell et al. 1995, Wright et al. 1994), giving rise to two loci DYA and DYB that are thought to have evolved from DQ (Ballingall et al. 2004a, Ballingall et al. 2004b). By contrast to other mammalian species, a high polymorphism of DRA genes is observed in ovine species (Ballingall et al. 2004b, Lewin et al. 1999, Wright and Ballingall 1994).

2.5- The MHC class I region

Three categories of MHC class I genes are described, the classical genes (class Ia), the non classical genes (class Ib) and the MHC related class I genes (MIC). We will mostly describe the features of MHC class Ia and Ib genes.

2.5.1- HLA class I region

The HLA class I region spans 1.9 Mb, from C6orf40 to MIC-B, and contains 6 expressed HLA class I genes (HLA-A, -B, -C, -E, -F and -G), 10 pseudogenes (HLA-H, -J, -K, -L, -P, -T, -U, -V and -W), and two MHC related class I genes, MIC-A and MIC-B. Among the six

expressed HLA genes, HLA-A, HLA-B, HLA-C are classical class I (Ia) genes and HLA-E, HLA-F and HLA-G non classical class I (Ib) genes.

2.5.2- SLA class I region

The SLA class I region spans 1.1 Mb (Renard et al. 2006). The SLA class Ia genes include three functional genes SLA-1, -2 and -3 and four pseudogenes, SLA-4, -5, -9 and -11. SLA-5 has all characteristics to be functional but no expression has yet been reported. Two MIC genes are known but only MIC-2 is predicted to be functional while MIC-1 appears to be a pseudogene.

By contrast to human for whom class Ia and Ib genes are intermingled into three clusters, all SLA class Ia genes map to a unique cluster between TRIM26 and TRIM39 and all class Ib and MIC genes map to another unique cluster in the most centromeric part of the locus (Figure II). These features are in agreement with a class I gene duplication that occurred after speciation in a species-specific manner leading to MHCs that harbour a common framework represented by non MHC molecules and histocompatibility molecules with a species-specific organization (Danchin et al. 2004).

Variation in the number of SLA-1 genes has been recently reported in swine and an additional gene termed SLA-12 has been characterized (Tanaka-Matsuda et al. 2009). In pig, there is a unique well sequenced reference haplotype, namely H01 or Hp1a.0 (Ho et al. 2009). The corresponding annotation is available online at the Vertebrate Genome Annotation (VEGA) database (<http://vega.sanger.ac.uk/index.html>). More haplotypes have to be fully sequenced and annotated to better describe all pig MHC features.

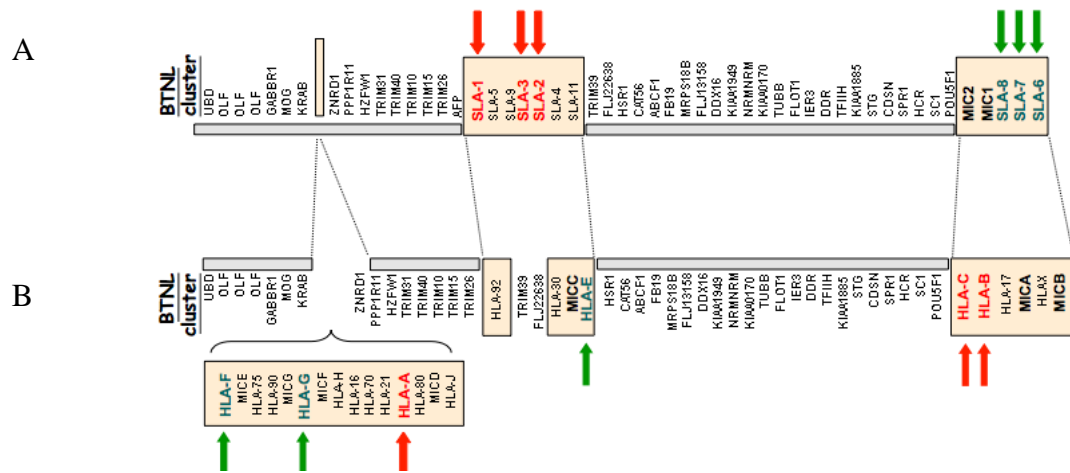


Figure II: The genomic organization of the pig and human MHC class I regions

A. The pig MHC class I region. B. The human MHC class I region. In both A and B the functional MHC class Ia genes are in red font and their position is indicated by red arrows. MHC class Ib genes are in green font and are marked by the green arrows. The orientation of the maps is from the telomere (left side) to the centromere (right side). The class I gene clusters are represented by boxes slightly shaded in pink. In human, the class I genes are distributed into three clusters whereas in swine, in only two clusters. The absence of MHC class I gene in the most telomeric cluster between KRAB and ZNRD1 is represented by an empty rectangle.

3- MHC class Ia genes and molecules

3.1- Organization and gene structure

MHC class Ia molecules are transmembrane glycoproteins of 44 Kilodaltons (KDa) that belong to the immunoglobulin superfamily. They are composed of an α chain that is non covalently associated to a β chain encoded by the β 2-microglobulin (β 2M) gene that maps outside the MHC region. At the cell surface, the membrane-bound MHC class I molecule is a trimolecular complex that includes the α chain, the β chain and the peptide, as shown in Figure III. These three components are prerequisite for the MHC complex cell surface expression.

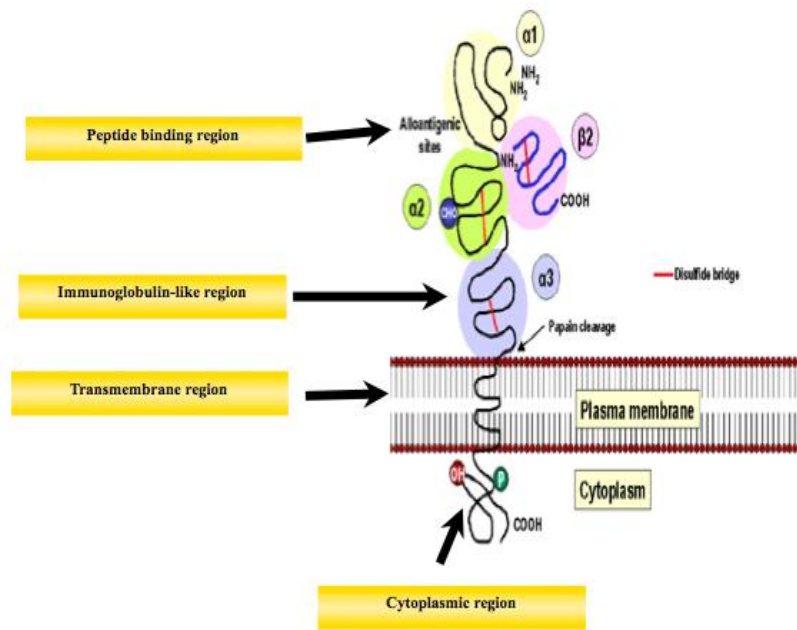


Figure III: The structure of MHC class I molecules

The MHC class I molecule is a heterodimer of a membrane spanning α chain that is bound non covalently to the $\beta 2$ -microglobulin. The $\alpha 1$ and $\alpha 2$ domains fold together to create a groove, which can bind a peptide.

Adapted from <http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=imm&part=A346>

The α chain of MHC class I molecules are encoded by MHC class I genes that have a common genomic organization in eight exons (Figure IV). Exon 1 corresponds to a leader peptide. Exons 2 and 3 encode the polymorphic $\alpha 1$ and $\alpha 2$ domains, respectively. Exon 4 corresponds to the $\alpha 3$ domain, which is an immunoglobulin-like region that binds to the T cell receptor. Exon 5 encodes the transmembrane domain and the remaining exons 6 through 8 stand for the cytoplasmic tail.

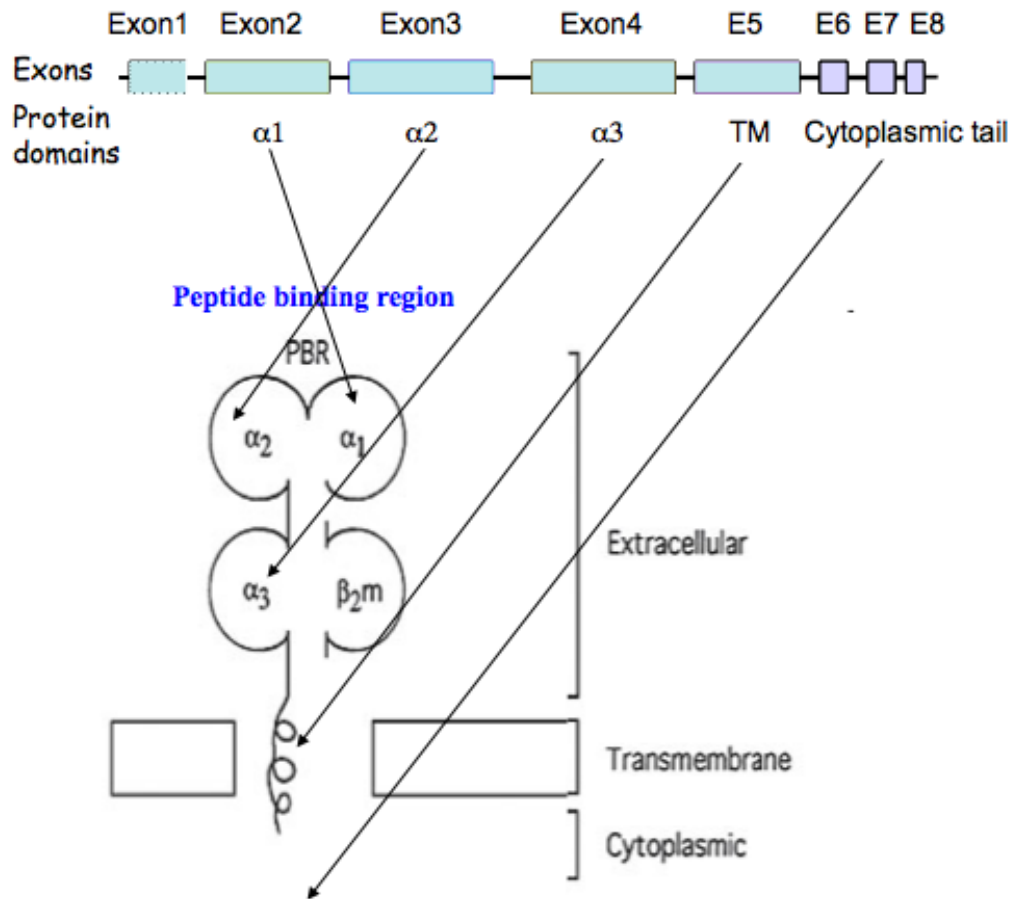


Figure IV: MHC class Ia gene structure

The $\alpha 1$ to $\alpha 3$ domains are encoded by exons 2 to 4, the transmembrane (TM) domain by exon 5, and the cytoplasmic tail by exons 6 to 8.

3.2- Expression

MHC class Ia genes are ubiquitously expressed in nucleated cells and expression levels may vary according to tissue types (Le Bouteiller 1994, Salter-Cid et al. 1998). The expression level of MHC Ia genes can be regulated by TNF, which can increase their expression

(Collins et al. 1986). In contrast, the infection of viruses, such as human cytomegaloviruses or adenoviruses can decrease the MHC Ia expression level (Miller et al. 1998, Trgovcich et al. 2006). Down-regulation of MHC Class I molecules has been described in several malignancies (Ritz et al. 2001). For example, in ovarian cancer, down-regulation of MHC Class I was associated with advanced stage of the disease and poor survival (Raffaghello et al. 2007, Rolland et al. 2007). In cervical cancer, only partial loss of MHC Class I was significantly associated with decreased overall survival (Badley and Frelinger 1991). The role of cancer immune escape mechanism in endometrial carcinoma is relatively unknown (Bijen et al. 2010). Modification of expression of MHC class Ia genes has also been reported during pregnancy in human (Trowsdale and Moffett 2008).

3.3- Functions

3.3.1- Peptide presentation to cytotoxic CD8+ T lymphocytes

MHC Class I molecules bind peptide antigens and display themselves to cytotoxic CD8+ T lymphocytes (CTLs). The expressed MHC antigen with non-self peptides are recognized as foreign antigens by T lymphocyte MHC restricted T cell receptors (TCR), which initiates the specific immune response (Germain and Margulies 1993). Thus, MHC class Ia molecules play a major role by providing immune surveillance against intracellular pathogens or viruses. The processing of MHC peptide presentation is a complex procedure that relies on many steps and components. In cells, cytoplasmic and newly synthesized proteins are degraded by the proteasome to 8–9 amino acid long peptides (Yewdell and Bennink 2001). The resulting peptides are translocated into the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP1 and TAP2) (Lankat-Buttgereit and

Tampe 2002). TAPs form the core of a multimeric peptide loading complex associating with the glycoprotein tapasin, which binds to nascent MHC class I heavy chain-B2M dimers. The loading complex also contains calreticulin, a lectin-like chaperone binding the monoglucosylated N-linked glycan on class I heavy chain (Radcliffe et al. 2002), and ERp57, a thioloxydo reductase noncovalently associated with calreticulin and disulfide-linked to tapasin (Dick et al. 2002). TAP-transported peptides, trimmed by a specific ER aminopeptidase (ERAP) (Saric et al. 2002, Serwold et al. 2002), bind TAP-associated MHC class I-B2M, inducing their release. An intact loading complex within the ER is critical for efficient MHC class I peptide association. Cells and mice lacking tapasin or TAP are deficient in MHC class I Ag presentation (Garbi et al. 2000). The molecular complex released from the ER comprises the class I molecule associated with the B2M chain. This peptide ligand is then driven through the Golgi apparatus for a glycosylation step for final suitability to be presented at the outer cell surface for T cell recognition (Figure V).

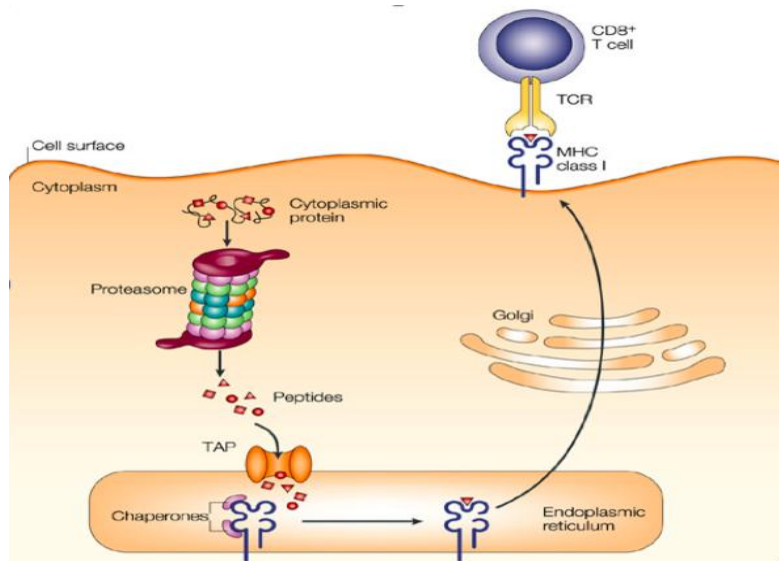


Figure V: MHC class I antigen presentation procedure

Cytosolic and nuclear proteins are degraded by the proteasome into peptides. The transporter for antigen processing (TAP) then translocates peptides into the endoplasmic reticulum (ER) while consuming ATP. MHC class I heterodimers wait in the ER for the third subunit, a peptide. Peptide binding is required for correct folding of MHC class I molecules and release from the ER and transport to the plasma membrane, where the peptide is presented to the immune system. TCR: T-cell receptor. Adapted from (Yewdell et al. 2003).

3.3.2- Modulation of Natural Killer (NK) cells

MHC class Ia molecules also regulate innate immunity as ligands for killer inhibitory receptors (KIRs) on NK cells. They bind to inhibitory receptors on NK cells that include *KIR* (killer cell Ig-like receptors) in man, C-type lectin-like Ly49 molecules in mouse, and *CD94/NKG2A* heterodimers in man and mouse (Anfossi et al. 2006, Uhrberg et al. 2001). Failure of MHC class Ia molecules to interact with these receptors may result in the killing of target cells as it occurs during tumour transformation or infection by certain viruses when target cells have lost or express insufficient amounts of MHC class Ia molecules. Each NK cell expresses at least one receptor specific for HLA class I molecules, while the co-expression of two or more self-reactive receptors is rare. This type of receptor distribution

allows the whole NK cell pool to detect the loss of even a single HLA class I allele on self cells, a frequent event in tumor transformation (Garrido et al. 1997). A common characteristic of the various HLA class I-specific inhibitory receptors is the presence, of immunoreceptor tyrosine-based inhibitory motifs in their cytoplasmic tail that enable them to recruit and activate SHP-1 and SHP-2 phosphatases (Lanier 1998, Long 1999, Moretta et al. 1996a). In turn, these phosphatases switch off the activating signalling cascade initiated by the various activating receptors. Provided that turning NK cells ‘off’ represents the major failsafe device to prevent the NK-mediated attack of normal HLA class I autologous cells, an ‘on’ signal must be generated upon interaction of NK cells with potential target cells. This signal is extinguished whenever appropriate interactions occur between inhibitory receptors and MHC class I molecules. On the other hand, the ‘on’ signal can be readily detected when NK cells interact with target cells that lack MHC class I molecules (Moretta et al. 1996b, Moretta and Moretta 2004, Moretta et al. 2004).

3.4- Polymorphism

One remarkable characteristic of MHC class Ia genes is their extremely high polymorphism, particularly in exons 2 and 3 that encode the $\alpha 1$ and $\alpha 2$ domains responsible for the peptide binding groove. This feature is not only a selective advantage for an individual to express class I molecules that bind different repertoires of peptides, but also for a population or species to have many class I variants segregating among its members (Bos and Waldman 2006). Given the strong linkage disequilibrium exhibited by the MHC loci, it is sometimes more appropriate and convenient for researchers to communicate and present findings in

terms of haplotypes (a specific combination of alleles of genes on the same chromosome) rather than individual allele specificities.

The human MHC Haplotype Project affirmed that HLA are the most polymorphic genes in the vertebrate genome with 300 total loci, including 122 gene loci with coding substitutions of which 97 were non-synonymous. In the HLA system, over several hundred alleles have been identified in HLA Ia. The three expressed human MHC classical genes HLA-A, -B and -C are highly polymorphic. Of the three classical class I genes, HLA-B is the most polymorphic with 1381 alleles known at this locus, and 960 for HLA-C. The amount of polymorphism within the peptide-binding region is extremely high, with 1001 HLA-A, 1605 HLA-B, and 690 HLA-C molecules characterised in human populations to date (IMGT/HLA database, <http://www.ebi.ac.uk/imgt/hla/stats.html>). The porcine MHC class Ia has fewer polymorphism than the human MHC class Ia molecules. To date, SLA-1 has 44 alleles matching to 44 proteins, SLA-2 has 46 alleles for 44 proteins, and 26 alleles corresponding to 26 proteins are reported for SLA-3 (<http://www.ebi.ac.uk/ipd/mhc/sla/stats.html>). Known SLA-Ia polymorphisms and organization into haplotypes are summarized in Table I.

Table I: SLA Ia polymorphism and haplotypes

Haplotype	Breed	Previous Designation	SLA-1	SLA-2	SLA-3
1a.0	Large White	H01	101	101	101
1b.0	Large White	H28	01rh28	101	01rh28
2	NIH, Sinclair, Hanford	a, b, H10	0201, 0701	201	null
3	NIH	c, H59	null	301	301
4a.0	NIH, Duroc	d, H04	401	401	401
4b.0	Yucatan	X	401	40201	401
4c.0	Meishan	K	401	401	401
5	Yucatan	W	401	w08sw01	05sw01
6	Yucatan	Y	08sy01	05sy01	601
7	Yucatan	Z	801	502	701
8	Westran	None	02we02, 04we01	07we01	302
9	Sinclair, Hanford	A	601	601	501
10	Sinclair	C	501	302	hm22
11	Sinclair	D	0101, w09sm09	501	0701sm19
12	Hanford	E	08sm08, w09sm09	10sm01	502
13	Hanford, Duroc	f, d2	w10sm21	w13sm20	401
14	Large White	H12	102	07rh12	01rh12
15	Large White	H34	102	05rh34	07rh34
16	Clawn	c1	401	w09an02	602
17	Clawn	c2	Blank	06an03	03an02
18	Meishan	M	401	06me01	304
19	Meishan	N	08ms05, 13ms21	w09sn01	602
20	Meishan	L	w10cs01, cs02	110102	101
21	Commercial breeds	H03	rh03	05rh03	601
25	Hampshire (LLC-PK1 porcine cell line)	None	1101	701	302
27	Duroc	d1	06an04, 08an03	102	101
56	Korean native pig	None	11jh01	jh01	303
59	Korean native pig	None	11jh02	jh02	503
60	Duroc	d2	an02	1002	502

4- MHC class Ib genes and molecules

MHC class Ib genes belong to the MHC class I gene series and present various features that led authors to classify them as non classical. These features include a very limited polymorphism, more restricted tissue expression, modification of the gene structure in the cytoplasmic tail, and expression of alternative RNA variants that encode putative protein isoforms.

4.1- Gene structure

MHC class Ib gene structure is usually modified in the number of exons that code for the cytoplasmic tail. Class Ia genes harbour eight exons with the last three exons specifying the cytoplasmic tail whereas the class Ib genes have either seven or eight exons. In the pig, the three SLA-Ib genes were initially shown to harbour seven exons but recent evidence has indicated that SLA-6 presents seven exons while SLA-7 and SLA-8 present eight exons.

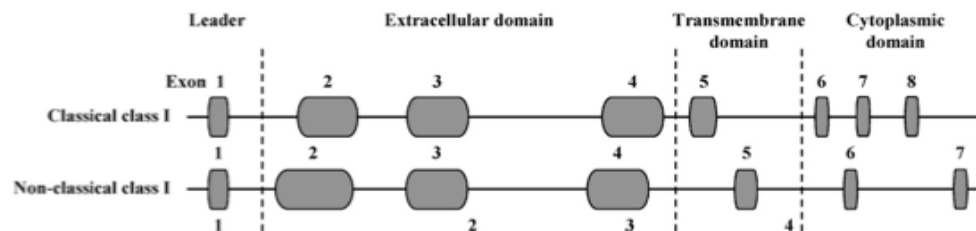


Figure VI: Comparative gene structure of SLA-Ia and Ib genes

In pig, class Ib genes were initially described with seven exons, the cytoplasmic tail being encoded by the last two exons instead of three exons for class Ia genes. Exons are represented by grey ovals and introns by lines (Renard et al. 2006).

4.2- HLA class Ib

The HLA complex contains three expressed class Ib genes, HLA-E, -F and -G. The HLA-Ib molecules were found to be specifically expressed mostly at the foeto- maternal interface (Ishitani et al. 2003) during pregnancy and were further intensively studied for their role in the control and/or maintenance of immunotolerance of the foetus by the mother. More generally, these molecules are found expressed at immunotolerant sites (Moscoso et al. 2006a, Moscoso et al. 2006b) and seem to play an important role in the maintenance of solid tumours (Le Maoult et al. 2004).

4.2.1- HLA-E

The HLA-E gene comprises seven exons and the full transcript HLA-E 001 contains 7 exons and encodes 358 amino acids as reported in the VEGA database (http://vega.sanger.ac.uk/Homo_sapiens/Info/Index) (Fig. VII). Orthology between HLA-E and the mouse gene H2-QaI has been demonstrated (Joly and Rouillon 2006).

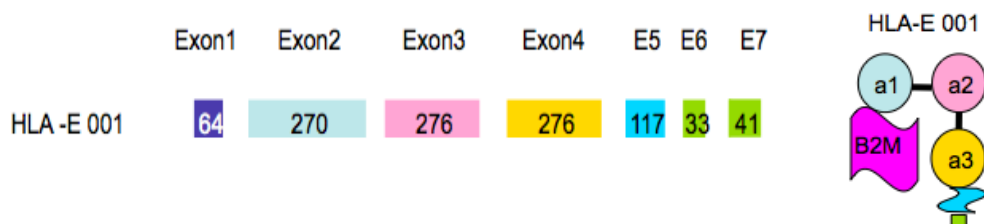


Figure VII: HLA-E transcripts and protein structure

Exons are represented by boxes and the colours correspond to the encoded protein domains. HLA-E protein structure is composed by all the parts necessary for cell surface expression.

RNAs are found in a wide range of tissues, as shown in Table II where the counts of HLA-E ESTs in different tissues are summarized, as provided by the NCBI EST Profile database (<http://www.ncbi.nlm.nih.gov/UniGene>). Although EST counts might not be a true indication of the protein activity, the expression patterns in these 44 tissues could be a good indicator. HLA-E EST was detected in 43 tissues except in ear. The highest count of transcripts per million (TPM) of HLA-E was recorded in adipose tissue (2289) followed by tonsil (1941), spleen (1445) parathyroid (1363), lung (816) and blood (753). The level of HLA-E expression can be augmented upon stimulation with interferon gamma (Koller et al. 1988, Mizuno et al. 1988).

HLA-E associates with B2M, interacts with NK cell receptors, and binds the $\alpha\beta$ CD8 T cell receptors (Rodgers and Cook 2005). HLA-E molecules interact with both inhibitory (CD94/NKG2A heterodimers) (Lee et al. 1998) and activating (CD94/NKG2C heterodimers) NK cell receptors (Rodgers and Cook 2005). HLA-E is expressed on cell surface and presents peptides derived from the leader sequence of MHC class Ia (O'Callaghan et al. 1998) and HLA-G molecules. The acquisition of these peptides by HLA-E appears to be tightly controlled and dependent on the expression of other MHC class I molecules, which serve as a source of these peptides together with functioning antigen processing machinery (Sullivan C. A. et al. 2009). HLA-E was also shown to present peptides derived from pathogens like the cytomegalovirus, thus providing evidence for a role in bridging innate and adaptive immune responses (Sullivan L. C. et al. 2008).

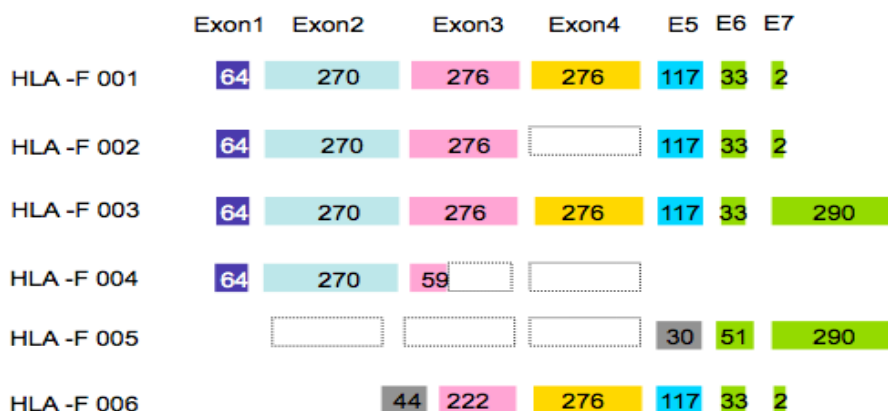
HLA-E is expressed at the foeto-maternal interface (Moscoso et al. 2006a) and its expression varies in tumour cells compared to normal cells. For instance, HLA-E is more expressed in melanoma cells than melanocytes in which no or very low HLA-E molecules are detected on

the cell surface. HLA-E expression on the cell surface in melanoma cells decreased their susceptibility to cytolytic activity by T cells, confirming a tolerogenic function (Derre et al. 2006).

4.2.2- HLA-F

The HLA-F gene structure includes seven exons and the full length transcript HLA-F-001 encodes a protein of 346 aminoacids (Figure VIII). Five HLA-F RNA variants are reported in the VEGA database that partly correspond to truncated RNAs and spliced variants leading to modified HLA-F isoforms that are different from the full length protein. There is no exon 4 in the HLA-F-002 transcript, therefore it is predicted to encode an isoform that lacks an $\alpha 3$ domain.

A



B

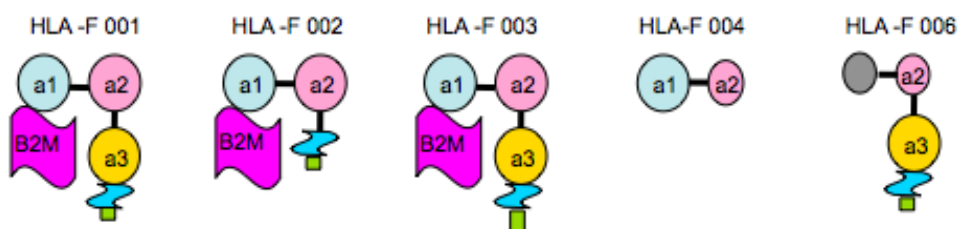


Figure VIII: HLA-F transcripts (A) and predicted protein structures (B)

Six HLA-F transcripts have been annotated in the VEGA database (http://vega.sanger.ac.uk/Homo_sapiens/Info/Index). Exons are represented by rectangles the colour of which corresponds to the protein domains. Missing exons due to alternative transcription are represented by empty rectangles.

HLA-F is mostly expressed in lymphoid tissues such as spleen (352 TMP), lymph node (283 TMP), tonsil (235 TMP) and thymus (184 TMP) as inferred from the EST database (Table II). HLA-F transcripts are also found in many other tissues, which indicate that they are widely expressed like HLA-E transcripts, but with a lower abundance.

HLA-F molecules can associate with B2M but seem to act independently on peptide presentation. The encoded molecules are predominantly expressed inside the cells (Apps et al. 2008, Boyle et al. 2006, Wainwright et al. 2000). Peptide presentation by HLA-F molecules has not been reported, but the possibility has not ruled out (Rodgers and Cook 2005).

Table II: Expression patterns of HLA-Ib genes inferred from EST sources

Tissue	# total EST	# HLA-E EST	TPM of HLA-E	# HLA-F EST	TPM of HLA-F	# HLA-G EST	TPM of HLA-G
Adipose tissue	13105	30	2289	1	76	0	0
Adrenal gland	33195	15	451	3	90	0	0
Gscites	40013	7	174	1	24	0	0
Bladder	29757	8	268	2	67	0	0
Blood	123476	93	753	14	113	0	0
Bone	71655	33	460	5	69	0	0
Bone marrow	48798	11	225	1	20	0	0
Brain	1100969	214	194	22	19	2	1
Cervix	48171	14	290	3	62	0	0
Connective tissue	149254	46	308	7	46	0	0
Ear	16212	0	0	0	0	0	0
Embryonic tissue	215722	11	50	0	0	0	0
Esophagus	20208	10	494	3	148	0	0
Eye	211052	42	199	11	52	0	0
Heart	89625	34	379	3	33	0	0
Intestine	234477	119	507	24	102	8	34
Kidney	211769	65	306	13	61	0	0
Larynx	24144	7	289	0	0	0	0
Liver	207739	30	144	4	19	0	0
Lung	336969	275	816	41	121	1	2
Lymph	44269	26	587	1	22	0	0
Lymph node	91607	35	382	26	283	0	0
Mammary gland	153267	88	574	1	6	0	0
Mouth	67053	13	193	3	44	0	0
Muscle	107711	34	315	4	37	0	0
Nerve	15768	8	507	0	0	0	0
Ovary	102050	45	440	13	127	0	0

Tissue	# total EST	# HLA-E EST	TPM of HLA-E	# HLA-F EST	TPM of HLA-F	# HLA-G EST	TPM of HLA-G
Pancreas	214811	109	507	14	65	2	9
Parathyroid	20540	28	1363	0	0	0	0
Pharynx	41328	15	362	0	0	0	0
Pituitary gland	16584	3	180	2	120	0	0
Placenta	280828	91	324	0	0	18	64
Prostate	189352	37	195	0	0	0	0
Salivary gland	20155	6	297	0	0	0	0
Skin	210574	113	536	5	23	0	0
Spleen	53953	78	1445	19	352	0	0
Stomach	96622	58	600	11	113	1	10
Testis	330449	24	72	3	9	4	12
Thymus	81130	29	357	15	184	1	12
Thyroid	47472	22	463	2	42	0	0
Tonsil	16999	33	1941	4	235	0	0
Trachea	52412	4	76	1	19	0	0
Umbilical cord	13680	3	219	0	0	0	0
Uterus	232876	74	317	15	64	0	0

*TMP: transcript per million

4.2.3- HLA-G

The HLA-G gene is reported with the most complex transcription pattern with a total of seven RNAs encoding seven protein isoforms (Figure IX). HLA-G encodes both membrane (HLA-G1 to G4) and soluble molecules (HLA-G5 to G7) which are generated by alternative RNA splicing (Carosella et al. 2003). HLA-G1 to G4 molecules are characterized by a short cytoplasmic tail encoded by exon 6. HLA-G2 lacks exon 3 that corresponds to the $\alpha 2$ domain; HLA-G3 lacks exons 3 and 4 and thus only has the $\alpha 1$ domain; HLA-G4 lacks exon 4 and hence the $\alpha 3$ domain. HLA-G5 and -G6 retain intron 4, which contains a stop codon that prevents the transcription of the transmembrane region and results in the expression of

soluble proteins. HLA-G7 retains intron 2 in which a stop codon results in the translation of a soluble protein that comprises only the $\alpha 1$ domain.

HLA-G is expressed in fetal extravillous trophoblast, adult thymic epithelial cells, cornea and nail matrix, meaning immune privileged sites (Ito et al. 2005, Le Discorde et al. 2003). As summarized in table II, among 44 tissues, HLA-G EST was only detected in 8 tissues (brain, lung, pancreas, stomach, testis, thymus, intestine, and placenta). HLA-G ESTs are specifically found in placenta (64 TPM) followed by intestine (34 TPM), thymus (12 TPM), testis (12 TPM) stomach (10 TPM) and pancreas (9 TPM). HLA-G is the most specifically expressed HLA-Ib gene with an abundance that is much lower than that of HLA-E and -F.

HLA-G inhibits cytolytic functions of NK cells and cytotoxic T lymphocytes and plays a key role in foeto-maternal tolerance during pregnancy (Hunt et al. 2007) and establishment of immune tolerance in tumorigenesis (Carosella et al. 2008).

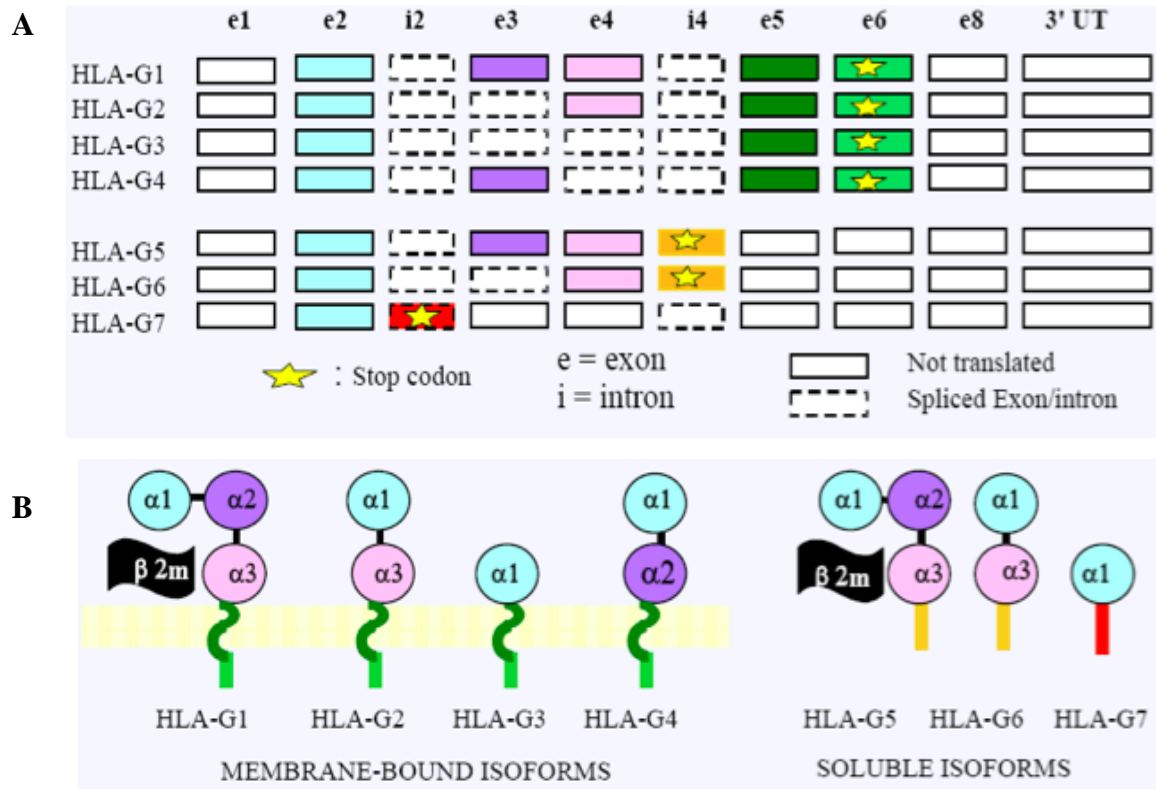


Figure IX: Alternatively HLA-G transcripts and protein isoform structure

Seven HLA-G transcripts encoding four membrane-bound and three soluble protein isoforms are reported (Carosella et al. 2003). B. HLA-G protein isoform.

Referred to <http://www.ensembl.org/>

The inhibitory effects of HLA-G are mediated through direct binding to inhibitory receptors ILT2, ILT4 and KIR2DL4, which are differentially expressed by immune cells (NK and some CD8+ T cells) (Ponte et al. 1999, Rajagopalan and Long 1999). Both full length membrane bound (HLA-G1) and soluble (HLA-G5) forms have been shown to have immunoregulatory functions, including the inhibition of T cell activation and stimulation of decidual NK cells and macrophages to produce cytokines that are beneficial to implantation (Le Bouteiller 2003, Le Bouteiller et al. 2003). HLA-G expression has been observed in

various malignancies: melanoma, ovarian, lung, cervical, colon, breast cancer, renal cancer and glioblastomas (Davies et al. 2001). Tumour occurrence and maintenance are frequently associated with a loss of HLA class Ia expression or abnormal expression of HLA class Ib antigen. Such peculiar HLA class I expression would allow tumour cells to escape not only from CD8+T, but also from NK-cell cytotoxicity. The association between the tumour and HLA-G is that HLA-G generates inhibitory signals in various immune cells that represent a mechanism used by tumour cells to escape from immunosurveillance (Garrido et al. 1997, Rouas-Freiss et al. 2005). A global scheme of HLA-G interactions with immune cells with functional implications is summarized in Figure X (Carosella et al. 2003). Functional homologies have been reported between HLA-G and H2-Qa2 (Comiskey et al. 2003).

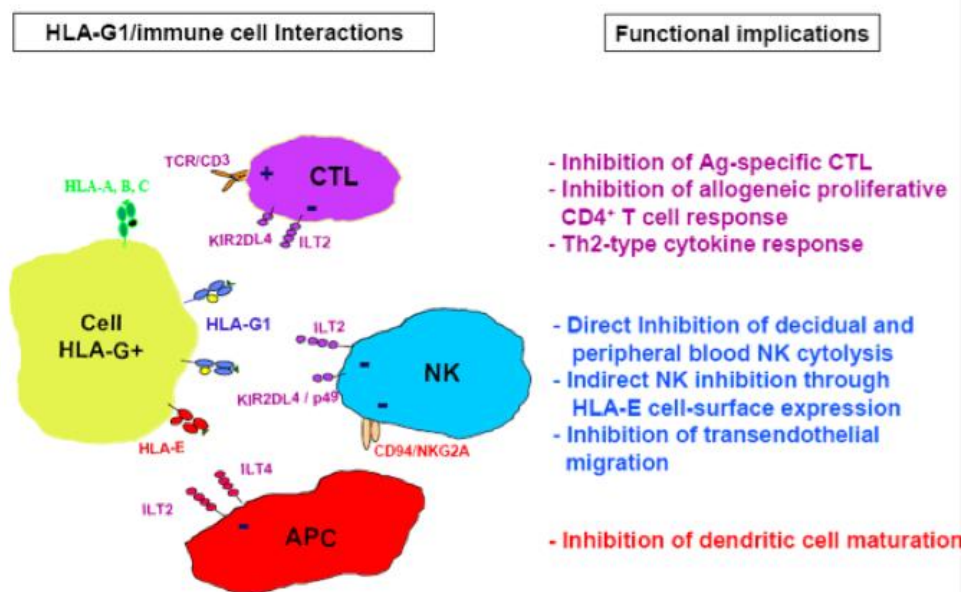


Figure X: HLA-G immune cell interaction

(Carosella et al. 2003)

4.3- SLA class Ib

The three genes SLA-6, -7 and -8 are classified as class Ib genes due to their limited polymorphism and slight variations in the three prime end specific of the cytoplasmic tail compared to the SLA-1, -2 and -3 genes (Chardon et al. 2001). The SLA-7 and SLA-8 genes were found to have a greater resemblance in coding regions to each other than to the SLA-6 gene (Chardon et al. 2001). No orthology or functional homology has ever been established with HLA class-Ib genes and neither gene mapping nor sequence phylogeny is helpful in this case (Chardon et al. 2001, Crew et al. 2004, Lunney et al. 2009, Renard et al. 2006).

It has been reported that SLA-Ib genes are transcribed in a less restricted manner than HLA class Ib genes but the number and variety of tissues included in this unique study were limited (Crew et al. 2004). Initially, SLA-6 has been referred to as PD6 and its expression was shown to be mostly restricted to secondary lymphoid tissues such as spleen and lymph node (Ehrlich et al. 1987). Expressions of the SLA-6 and SLA-8 mRNA transcripts have been detected in a variety of tissues with very low levels in the brain. SLA-7 transcripts exhibited more limited tissue distribution with high levels in thymus, and none detected in the kidney, brain and peripheral blood mononuclear cells. SLA-8 expression was likewise ubiquitous except no transcripts were detected in brain. The highest levels were observed in thymus, but no transcripts were detected in kidney or PBMC, SLA-7 transcripts were barely perceptible in spleen and testes (Crew et al. 2004). Evidence suggested that the SLA-6 gene may undergo alternative splicing, similar to the non-classical HLA-Ib gene (Lunney et al. 2009).

Sequence alignments revealed that SLA class Ia and Ib promoter regions contain a series of conserved putative regulatory motifs located within distal and proximal promoter domains.

An interesting study was carried out in order to test the constitutive and inducible activity of SLA-Ia and Ib promoters (Tennant et al. 2007). The promoters of SLA-1, -2, -3, -6 and -7 were cloned upstream from the luciferase reporter gene and the recombinant expression vectors were transiently transfected into Max cells, immortalized pig cells, that respond to interferon and TNF- α . Both classical and non-classical promoters were constitutively active. By contrast to SLA-Ia promoters, SLA-7 and -6 promoters did not respond to interferon alpha or gamma. This was confirmed by the transactivation of SLA-1, but not SLA-7, after the co expression with interferon regulatory factors (IRFs), IRF-1, IRF-2, IRF-3, IRF-7, and IRF-9. The response of SLA-1 and SLA-7 to the pro-inflammatory cytokine TNF- α was also investigated. The TNF- α treatment resulted in a twofold induction of the SLA-1 promoter but a smaller induction for the SLA-7 promoter. These results suggest distinct regulatory systems for pig MHC class Ia and Ib genes, as seen in human MHC the authors conclude on the importance of regulation variations in antigen presentation during infection (Tennant et al. 2007).

The three SLA-Ib genes are predicted to encode membrane-anchored glycoproteins that could associate with B2M and bind peptides (Chardon et al. 2001), but these features have never been demonstrated.

4.4- Polymorphism

Based on the IPD-MHC SLA database updated May 2008, SLA-Ib genes are less polymorphic than SLA Ia genes. There are only 9 SLA-6 alleles and 2 alleles for both SLA-7 and SLA-8. Conversely, 116 alleles have been identified for SLA Ia alleles: 44 SLA-1

alleles, 46 SLA-2 alleles, and 26 SLA-3 alleles. The extreme polymorphisms of the SLA Ia genes are, as expected, concentrated in exon 2 and 3 of the coding regions. Polymorphisms also include copy number variations (CNVs). CNVs are reported for non classical genes in rat and cattle. One to four MHC class Ib genes have been identified in rat according to haplotypes (Lau et al. 2003) and four MHC class Ib genes have been characterized in cattle (Birch et al. 2008). As in swine, The HLA-E, -F, and -G genes exhibit very low levels of allelic polymorphism. These low levels of allelic polymorphism presumably reflect their respective specialized functions.

Table III: The comparison of polymorphism between class Ia and class Ib genes

Human Gene	No Poly.	Porcine Gene	No Poly.
HLA-A	1381	SLA-1	44
HLA-B	1927	SLA-2	46
HLA-C	960	SLA-3	26
HLA-E	9	SLA-6	9
HLA-F	21	SLA-7	2
HLA-G	46	SLA-8	2

In swine, SLA-Ib polymorphisms are usually not studied in haplotype characterization. In table IV, scarce information on SLA-6 alleles and known SLA haplotypes are summarized.

Table IV: SLA-6 polymorphism and SLA haplotypes

Haplotype	Breed	Previous designation	SLA-6 allele
1a.0	Large White	H01	101
1b.0	Large White	H28	ND
2	NIH, Sinclair, Hanford	a, b, H10	w02sa01
3	NIH	c, H59	103
4a.0	NIH, Duroc	D, H04	102
4b.0	Yucatan	X	104
4c.0	Meishan	K	104
5	Yucatan	W	Null
6	Yucatan	Y	03sy01
7	Yucatan	Z	101
8	Westran	None	01we01
9	Sinclair, Hanford	A	ND
10	Sinclair	C	ND
11	Sinclair	D	ND
12	Hanford	E	ND
13	Hanford, Duroc	F, d2	ND
14	Large White	H12	ND
15	Large White	H34	ND
16	Clawn	c1	ND
17	Clawn	c2	ND
18	Meishan	M	102
19	Meishan	N	105
20	Meishan	L	103
21	Commercial breeds	H03	ND
25	Hampshire (LLC-PK1 porcine cell line)	None	ND
27	Duroc	d1	ND
56	Korean native pig	None	w04jh01
59	Korean native pig	None	102
60	Duroc	d2	ND

5- Aim of the work

Intensive research is carried out on HLA-Ib genes but very limited information is still available in swine. HLA-Ib molecules interact with NK-cell receptors and are involved in immunomodulation, allergy, autoimmunity and embryonic development. They also bind to T cell receptors and bridge innate and adaptive immunity. Indeed, SLA-Ib genes are highly interesting candidate genes for immune-tolerance functions in swine and functional studies are highly desirable to tests such hypotheses. SLA-Ib molecules are predicted to have the ability to present peptides on cell surfaces, but the proteins have never been detected due to the lack of specific antibodies. Very scarce information is available on tissue specificity and the splicing transcript patterns need to be described in order to know whether various isoforms exist for a unique gene as reported for HLA-G. HLA-E, -F and -G molecules probably play distinct and complementary functions as likely expected for SLA-6, -7 and -8.

Our aim was to study SLA-Ib genes at the DNA, RNA and protein levels. The work was further divided into three major parts related to transcription studies, protein expression and polymorphism characterization. We have studied the transcription levels of the three genes SLA-6, -7 and -8 in a wide range of adult pig tissues and have described the splice variants of each gene. Monoclonal antibodies were produced by a private company and transfection experiments with recombinant vectors suitable for expression of full length proteins were designed to specifically detect the proteins. Polymorphism studies were carried out by amplification and further sequencing of the full length genes from upstream the promoter to downstream of the polyadenylation signal in order to characterize allelic variations within coding and regulatory sequences. For our research purpose, we used biological samples from two pig breeds, French Large White and MeLiM pigs. Large White pigs are highly selected

for production traits. MeLiM pigs are minipigs bred in the laboratory, which present a genetic susceptibility to melanomas.

Materials and Methods

1- Transcription studies

1.1- Animals and samples

Two pig breeds were included in our study, French Large White and Melanoma bearing Llibechov Minipigs (MeLiM) pigs. French Large White pigs correspond to commercial animals highly selected for meat production and quality and bred in an experimental farm unit. MeLiM animals belong to an inbred closed line selected for the spontaneous occurrence and regression of cutaneous melanomas (Horak et al. 1999, Vincent-Naulleau et al. 2004). These animals are maintained in our experimental facilities for biomedical research with the aims to identify the genetic factors that control the disease susceptibility and to decipher the mechanisms involved in the tumour regression. Adult tissues were sampled from two Large White and three MeLiM pigs. We have also used primary fibroblasts derived from a Large White pig that harbours the Hp1a.0 haplotype corresponding to the reference SLA haplotype in pig (Renard et al. 2006). Tissues were frozen in liquid nitrogen and stored at -80°C before RNA extraction.

1.2- RNA extraction

Total RNAs had been extracted by Steliana Kevorkian during her master degree in the laboratory before I arrived to the laboratory (Kevorkian, 2007). Various RNA extraction kits (Qiagen, France) were used according to tissue characteristics. The RNeasy Midi kit was used for duodenum, ileum, jejunum, kidney, liver, lung, Peyer's patches, spleen, thymus, tonsil, cornea, epididymis, testis, ovary, uterus, adrenal gland, inguinal lymph node as well

as pig fibroblasts. The RNeasy Fibrous Tissue kit was chosen for diaphragma, longissimus dorsi, heart (ventricle), nasal mucosa, and aorta. The RNeasy Lipid Tissue kit was used for back fat, skin and brain. RNA extractions were carried out following the manufacturer's instructions. For all samples, residual genomic DNA was removed by on-column digestion of DNA with DNase I (DNase set and cleanup, Qiagen, France). RNA concentration was determined by Nanodrop quantification (Thermo Fisher Scientific Inc., USA). RNA quality was checked on an Agilent 2100 Bioanalyzer (Agilent Technologies, Germany). RNAs with a RIN score between 8 and 10 were used. The concentration of RNA stock solutions was adjusted at 1 µg/µl in water for storage at -80°C.

1.3- Primer design

All primers were designed using the Primer3 program available online at <http://frodo.wi.mit.edu/primer3/>. For the three SLA-Ib genes, primers targeting coding sequences were designed according to reference mRNA sequences (SLA-6: NM_001113704; SLA-7: NM_213768; SLA-8: NM_001113703) (Crew et al. 2004) and primers specific of the three prime untranslated regions (3'UTRs) were designed according to the sequence of the BAC clone that comprises the three SLA-Ib gene (AJ251914) (Chardon et al. 2001). Primers were designed in order to amplify either partial coding sequences from exon 4 to 3'UTRs or full coding sequences from exon 1 to 3'UTRs. For SLA-7, a primer set was also designed to amplify genomic DNA from intron 6 to 3'UTR in order to check that a splice variant was not due to a deletion in the corresponding genomic DNA. All primers were used matching the sequences from exon1 to 3'UTR. The gene

RPL32 was chosen as a reference gene for RNA expression levels. All primers are summarized in table V.

Table V: Gene specific primers used for transcription studies

Gene	Primer name	Primer sequence	Position	Reference number	Expected size*
SLA-6	SLA6_e1_F	GGAAGGATGCAGGTCACG	Exon 1	NM_001113704 AJ251914.1	1221bp
	SLA6_3UTR_R	GCAAGGCAGACACATTCAGA	3'UTR		
	SLA6_e4_F	CTTCTGGAGAGGAGCAGAGC	Exon 4		813bp
	SLA6_3UTR_R	GCAAGGCAGACACATTCAGA	3'UTR		
	SLA6_e2_F	GGCCACGGTAGTGACCTTTA	Exon 2		1369bp
	SLA6_e4_R	ACTCTAAGATGCTGGGCCCT	Exon 4		
SLA-7	SLA7_e1_F	ATGGGGCCCCGAGCCCTCCTCCT	Exon 1	NM_213768 AJ251914.1	1465bp
	SLA7_3UTR_R	AGAGCCACTGCTGATCCAGT	3'UTR		
	SLA7_e4_F	TGGAGAGGAGCAGAGCTACA	Exon 4		687bp
	SLA7_3UTR_R	AGAGCCACTGCTGATCCAGT	3'UTR		
	SLA7_i6_F*	GCTGAGATCCCCAAAACCTT	Intron 6		628bp
	SLA7_3UTR_R	AGAGCCACTGCTGATCCAGT	3'UTR		
SLA-8	SLA8_e1	ATGGAGTCTCAGATGCTTCTTC	Exon 1	NM_001113703 AJ251914.1	1148bp
	SLA8_3UTR_R	TGGAGAAGCCAGTCTTCCAT	3'UTR		
	SLA8_e4	CCTGGAGAGGAGCAGAGCTA	Exon 4		957bp
	SLA8_3UTR_R	TGGAGAAGCCAGTCTTCCAT	3'UTR		
RPL32	RPL32_F	TGCTCTCAGACCCCTTGTGAAG	Exon 1	NM_001001636	320bp
	RPL32_R	TTTCCGCCAGTTCGCTTA	Exon 4		
B2M	B2M_F	ACTTTTCACACCGCTCCAGT	5'UTR	NM_213978.1	430bp
	B2M_R	GGATTCATCCAACCCAGATG	Exon 4		

* Expected size: Indicates the longest PCR product expected by referring to the reference sequence.

1.4- Reverse Transcription

Two and half μg of DNaseI-treated total RNA were reverse-transcribed using Superscript II enzyme with Oligo dT₁₂₋₁₈ (Invitrogen, Carlsbad, CA) and random primers (Promega, USA) as recommended by the manufacturer (Invitrogen, Carlsbad, CA). Reverse Transcription (RT) reactions were carried out in a final volume of 30 μl with RNase OUT (20 units). RT reactions were adjusted to 50 μl with water and further considered as the non diluted RT products. Random Primers were used to prime mRNAs with or without poly A for cDNA synthesis. The RNase OUT, which is a recombinant ribonuclease inhibitor, was added to the reaction as a potent non-competitive inhibitor of pancreatic-type ribonucleases such as RNase A.

1.5- PCR amplification

SLA Ib cDNAs were amplified with the primers spanning exon1 to the 3'UTR or exon 4 to the 3'UTR (Table V). PCR reactions were carried out in 15 μl with 1 μl of a fivefold dilution of the RT products. The following PCR cycling parameters were used: denaturation at 94°C for 3 min, followed by 35 cycles of amplification (94°C denaturation for 1 min, around 60°C annealing for 30 s, and 72°C extension for 1 min 30 s) and 5 min for a final extension at 72°C. The PCR products were run on a 1.5% agarose gel with ethidium bromide and the separated DNA fragments were viewed under a UV light source. For the further cloning and sequencing, PCR products were excised from gels and isolated as described in the following section (DNA fragment purification).

1.6- DNA fragment purification

1.6.1- Purification of DNA from agarose gels

For SLA Ib PCR products with more than one band, all fragments were sliced from the agarose gel and purified with the QIAquick Gel Extraction Kit (QIAGEN, USA). The QIAquick Gel Extraction was chosen because it is suitable for fast cleanup of up to 10µg of DNA fragments from enzymatic reactions and agarose gels. The protocol was as follows: excise the DNA fragment from the agarose gel, weigh the gel slice in a colourless tube; add 3 volumes of Buffer QG to 1 volume of gel (100 mg or approximately 100 µl); Incubate in 50°C water bath for 10 min or until the gel slice to completely dissolved; After the gel slice has dissolved completely, add 1 gel volume of isopropanol to the sample and mix; Place a QIAquick spin column in a provided 2 ml collection tube; To bind DNA, apply the sample to the QIAquick column; Wash with wash buffer QIAquick column; Elute DNA with Buffer EB (10 mM Tris·Cl, pH 8.5). The cleaned-up DNA was quantified by Nanodrop quantification (Thermo Fisher Scientific Inc., USA) and visualized by running a 1% agarose gel electrophoresis.

1.6.2- Direct purification from PCR

For the PCR products with only one band, the amplified DNA fragments were directly purified from PCR with the QIAquick PCR purification kit according to the manufacturer instructions (QIAGEN, USA). This kit was chosen because it is suitable for fast cleanup of PCR products over 100 bp from PCR reaction directly. DNA was eluted in 50µl of water or elution buffer (10 mM Tris-HCl, pH 8.5) provided by the kit. The purified DNA was

quantified by Nanodrop (Thermo Fisher Scientific Inc., USA) and visualized by running a 1% agarose gel electrophoresis.

1.7- PCR fragment cloning and sequencing

1.7.1- Ligation reaction

The PCR products were inserted into pCR® 2.1 vector using the TA Cloning® Kit (Invitrogen, USA) designed for cloning purified TAQ-amplified PCR products that are smaller than 4kb. The ratio of insert: vector was 1:3. Ligation reactions were carried out using 50 ng of linearised dephosphorylated vector, purified DNA in order to respect the 1:3 ratio, 1ul of a 10X ligation buffer, 1-2 U of T4 DNA ligase and water up to a final volume of 10ul. Ligation reactions were performed overnight at 14°C. To calculate the appropriate amount of insert DNA segment to be included in the ligation reaction, the following equation was used:

$$X \text{ ng PCR product} = \frac{(Y \text{ bp PCR product})(50 \text{ ng pCR}^{\circledR} 2.1 \text{ vector})}{(\text{size in bp of the pCR}^{\circledR} 2.1 \text{ vector: } \sim 3900)}$$

1.7.2- Preparation of electro-competent bacteria

DH10B *E coli* bacteria were inoculated into 50 ml of LB medium (without antibiotics) from fresh colonies and grown overnight at 37°C with shaking. The next day, 5ml of this pre-culture were transferred into 500 ml of LB medium (without antibiotics) and grown for 3 to 5 hours with vigorous shaking at 37°C until the OD₅₅₀ reaches 0.6-0.7. Cells were collected by centrifugation at 4,000 rpm (centrifuge J21) at 4°C for 10 min and the pellets were resuspended in 500 ml ice-cold sterile water. Cells were centrifuged again and the pellets

resuspended in 500 ml ice-cold sterile water with 10% glycerol. Cells were centrifuged again, the pellets resuspended in 20 ml of ice-cold water with 10% glycerol and the cell suspension was transferred into 50 ml plastic tubes for a final centrifugation (4,700 rpm, centrifuge Eppendorf 5810R). Pellets were finally resuspended in 500µl of ice-cold sterile water with 10% glycerol and the volume was adjusted in order to have a cell suspension with an $OD_{550nm}=140$ that corresponds to an $OD_{550nm} = 0.7$ for 5 µl of cell suspension diluted into 1 ml of water. Ready-to-use electro-competent bacteria were aliquoted by 30µl in 1.5 ml microtubes and stored at -80°C before use.

1.7.3- Transformation of competent bacteria

One ul of each ligation reaction were added to 50 ul of frozen electro-competent DH10B bacteria thawed on ice and transferred into electroporation cuvettes. Electroporation was performed under the following conditions: 2500 V, 25 µF, 201 Ω, 5ms (Electroporator Eppendorf 2510). After electroporation, 100 µl of LB medium was added to the cell suspension containing DNA and transferred into a 15-ml culture tube containing 900 µl LB medium without antibiotics. The cell suspension was incubated for 45 to 60 min at 37°C with shaking 250rpm. 200 ul of the cell culture were plated on LB agar plates supplemented with 100 µg/ml ampicillin as well as 200 µg/ml IPTG and 200 µg/ml X-Gal and for white/blue selection. The plates were incubated overnight at 37°C.

1.7.4- Screening of recombinant bacteria

Recombinant bacteria were screened by PCR using crude bacteria as DNA templates. PCRs were carried out in 15 µl using vector primers present on both sides of the cloning site (5'-GCGGATAACAATTTTCACACAGG-3' and 5'-TGTAACGACGGCCAGTGAATTG-3'). Inserts were sized by visualizing PCR products separated by electrophoresis in agarose gels. Recombinant bacteria with inserts having expected sized were selected for DNA preparation, sequencing and long term storage in 80% LB medium supplemented with ampicillin and 20% glycerol at -80°C.

1.7.5- Plasmid DNA preparation

Pure plasmid DNA was prepared from 5 ml bacterial culture using the S.N.A.P. Miniprep kit (Invitrogen, USA) that is suitable to quick and efficient extraction of highly pure plasmid DNA ready for automated sequencing.

To confirm the insert size, the purified plasmid DNA was digested by the restriction enzyme EcoRI. The digestion reactions were set up as follows: 100~300ng (~2µl) of plasmid DNA, 1.5 µl 10X digestion buffer, 1µl (2U) EcoRI and water up to a final reaction volume of 15 µl. The reaction mixture was incubated at 37°C for 2 hours. After incubation, the digested DNA fragments were separated by electrophoresis in a 1.5% agarose gel.

1.7.6- DNA sequencing

The clone inserts were sequenced with the forward and reverse primers specific for the vector and the primers specific for the SLA Ib sequences (see table VI). 15µl of plasmidic

DNA solution at 75ng to 150ng/ul were sent for Sanger sequencing by the MWG Company (<http://www.mwg-biotech.com/>).

Table VI: Universal and internal primers used for SLA Ib transcript sequencing

Primer	Sequences (5' to 3')
SLA-6 exon4-F	CTTCTGGAGAGGAGCAGAGC
SLA-6 3'UTR-R	ACTCTAAGATGCTGGGCCCT
SLA-7 exon4-F:	TGGAGAGGAGCAGAGCTACA
SLA-7 exon4-R:	CTAAGGAGATCTCCCTGACC
SLA-8 exon4-F	CCTGGAGAGGAGCAGAGCTA
SLA-8 exon4-R	TCACCCTGAGATGGGGTAAG
M13 Reverse primer	CAGGAAACAGCTAGTAC
T3 primer	TAATACGACTCACTATAGG

1.8- Sequence analysis

Sequence analyses were carried out in order to check the quality of the sequences and make comparison with existing genomic and transcript sequences. The chromatograms were visualized with the software Chromas to check sequence quality. Sequence alignments were carried out with the blast alignment search tools available at NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Multiple alignments were carried out with CLUSTAL W (Thomson et al., 1994) available at <http://align.genome.jp/>. cDNA sequences were translated to aminoacids by the online DNA to Protein translation tool (<http://bio.lundberg.gu.se/edu/translat.html>).

2- Protein studies

2.1- Protein structure prediction

The protein structure prediction was done with CPHmodels 3.0 (<http://www.cbs.dtu.dk/services/CPHmodels/>), which is a protein homology modeling server based on profile-profile alignment guided by secondary structure and exposure predictions (Pedersen et al. 2008).

2.2- Construction of expression vectors

2.2.1- Cloned cDNAs

Five cDNAs specific to the full coding sequences of SLA-6, SLA-7, SLA-8, SLA-1 and B2M from the ATG to the stop codon were prepared for cloning into expression vectors. Main features of the cDNAs are summarized in table VII and the corresponding nucleotide and amino acid sequences are in supplementary files S1 and S2.

Table VII: cDNA sequence features

gene	Animal*	Tissue	# clone	Size (bp)	ATG	TGA	polyA	# exons
SLA-1	M_485	Thymus	18	1387	yes	yes	no	8
SLA-6	M_485	Thymus	6	1221	yes	yes	no	7
SLA-7	M_485	Heart	14	1465	yes	yes	no	8
SLA-8	M_485	thymus	6	1123	yes	yes	no	8
B2M	M_485	thymus	1	430	yes	yes	no	4 (NCBI)

*M: MeLiM pig

2.2.2- Vectors for transfection in mammalian cells

The cDNAs specific for SLA-6, -7 and -8 were cloned into the pVAX1 vector (cloning site HindIII/XhoI) under the cytomegalovirus (CMV) promoter suitable for strong expression of recombinant proteins in mammalian cells upon transient and stable transfection. The vectors were constructed by In Cell Art Company (Nantes, France) in the frame of the production of SLA-Ib specific antibodies. The three constructed vectors referred to as pVAX1-SLA6, pVAX1-SLA7 and pVAX1-SLA8 have the same restriction map (Figure XI).

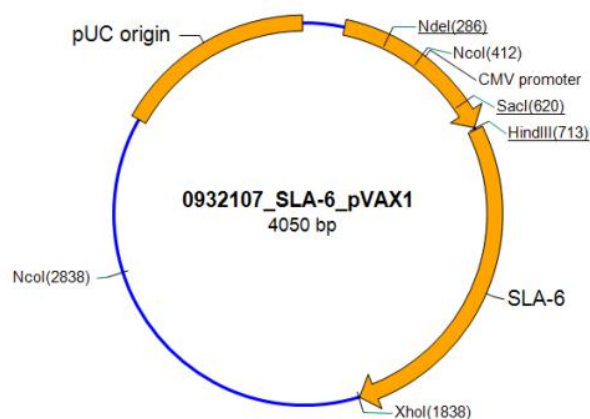


Figure XI: Restriction map of the plasmid pVAX1-SLA6

2.2.3- Vectors for transfection in drosophila cells

The five cDNAs were each cloned into the pAc5.1/V5-His vector (Invitrogen, USA) under the promoter of the drosophila Actin 5C gene (Figure XII). This 5.4 kb long vector is suitable for strong expression of recombinant proteins in drosophila cells either transiently or stably transfected. The vector pAc5.1/V5-His vector was digested by EcoRI and dephosphorylated with Calf Intestinal Alkaline Phosphatase to remove the 5'Gp. The

sequences to ligate were removed from the pCR® 2.1 vector by a digestion with the restriction enzyme EcoRI that is on both sides of the TA cloning site. Inserts were ligated to the linearized vectors and the ligated products transformed into DH10B electrocompetent bacteria as previously described in the section dedicated to transcription studies. The general scheme of the vector construction is presented at Figure XIII. For stable transfection, drosophila cells can co-transfected with a selection vector that is either pCoHygro for selection with hygromycin or pCoBlast for selection with blasticidin.

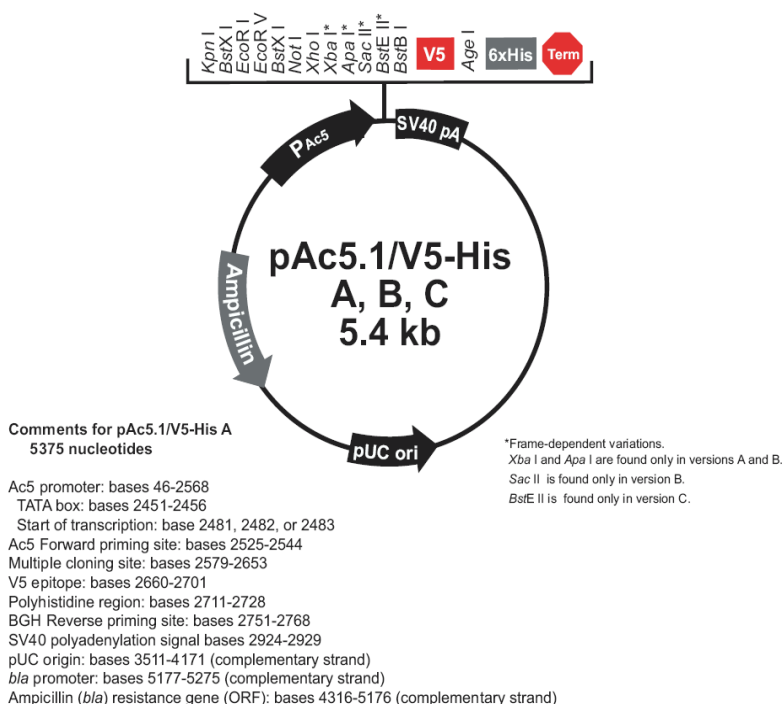


Figure XII: Map and features of the expression vector pAc5.1/V5-His

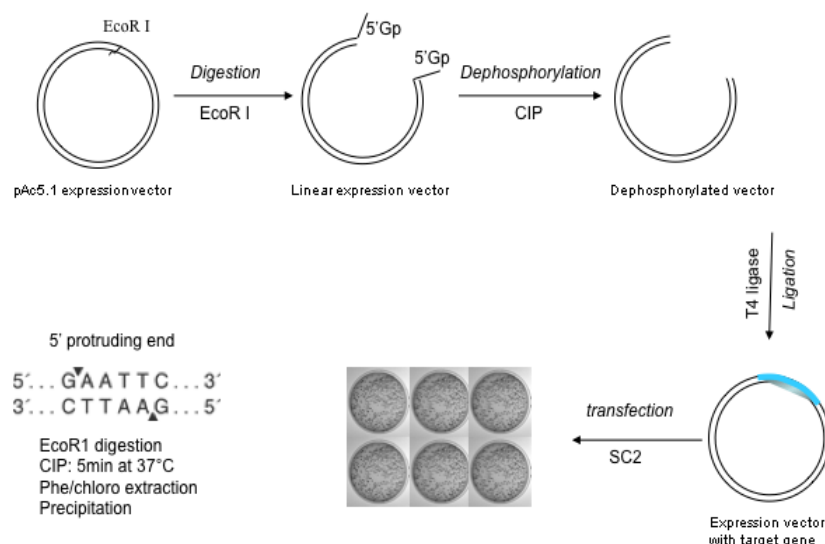


Figure XIII: Cloning of cDNAs into the expression vector (pAc5.1/V5-His vector)

2.3- Large scale preparation of plasmid DNA

Large amounts of highly purified plasmids are required for efficient transfections. Large scale plasmid DNA batches were prepared using the PureYield Plasmid Maxiprep System (Promega, USA). The procedure was as follows: recombinant clones were grown in 100 to 250 ml of LB medium supplemented with 100 µg/ml ampicillin overnight (16–21 hours). The cells were collected by centrifugation at $5,000 \times g$ for 10 minutes at room temperature. The cell pellets were resuspended in 12 ml of Cell Resuspension Solution, followed by lysing bacterial cells in 12 ml of Cell Lysis Solution, and mixing gently. After incubating lysed bacterial cells for 3 minutes at room temperature, 12 ml of Neutralization Solution were added to the lysed cells and the solution was mixed by gently inverting the tube 10 to 15 times. The lysate was collected by centrifugation at $14,000 \times g$ for 20 minutes at room

temperature in a fixed-angle rotor. After spinning, the cleared lysate was poured onto a blue PureYield™ Clearing Column on the top of a white PureYield™ Maxi Binding Column, which was placed onto the vacuum manifold in order for the plasmid DNA to bind to the column membrane when the lysate passed through columns. Followed the further washing, the binding membrane was dried by applying a vacuum for 5 minutes. To elute the DNA bound on the membrane, 1ml of Nuclease-Free Water was added onto the binding column. The DNA solution was collected using the Eluator™ Device. The DNA quantity and quality were checked by Nanodrop. The purified DNA was checked by agarose gel electrophoresis.

2.4- Cell lines and transfection

2.4.1- Insect cells: the drosophila Schneider 2 cells (SC2)

2.4.1.a- Cell culture conditions

SC2 cells are Drosophila semi-adherent cells that grow at 28°C or room temperature without CO₂. The cell line was provided by the Laboratory of Virology and Molecular Immunology (INRA, Jouy-en-Josas, France). These cells are highly interesting to study expression of MHC molecules on the cell membrane because there is an accumulation of the molecules on the cell surface at low temperature instead of a back and forth traffic of MHC molecules between the surface and the cytoplasm in mammalian cells that grow at 37°C. Moreover, there is no basal level of MHC molecule expression in SC2 cells. These cells have been shown to be very relevant to set up functional studies on MHC molecules (US Patent 6255073 - Antigen presenting system and methods for activation of T-cells).

Cells are grown in the Schneider Drosophila Medium (Invitrogen, USA) supplemented by 10% heat-inactivated fetal bovine serum (FBS; Invitrogen Catalog no. 16000-044) and 1% Penicillin-Streptomycin (Invitrogen Catalog no. 15070-063).

Cells were seeded at a density of 5×10^5 cells/ml in flasks with loosen caps for oxygenation/aeration. Cells were incubated a 28°C in a non-humidified, ambient air-regulated incubator or at room temperature. When the culture density reaches 6 to 20×10^6 cells/ml, the cells can be used for transfection, reseeded in other plates or frozen. Cells were harvested without trypsinization and centrifuged before resuspension in the appropriate solution or culture medium.

2.4.1.b- Transient transfection of SC2 cells

3×10^6 SC2 cells were plated in 35mm dishes with 3ml complete Schneider's Drosophila Medium to get a density of 1×10^6 SC2/ml. Cells were grown at 28°C for 6 to 16 hours until they reach a density of $2-4 \times 10^6$ SC2/ml. Transfections were carried by a calcium/phosphate precipitation method using the Drosophila Expression System kit (Invitrogen, Carlsbad, CA). For one 35 mm plate, the solutions were prepared as follows:

<u>Solution A:</u>	2 M CaCl_2	36 μl
	Recombinant DNA (19 μg)	X μl
	Tissue culture sterile water	300 μl

Solution B: 300 μl 2X HEPES-Buffered Saline (50 mM HEPES, 1.5 mM Na_2HPO_4 , 280 mM NaCl, pH 7.1.

Solution A was slowly added solution B drop after drop with continuous mixing. The mixture was incubated at room temperature for 30-40 minutes and drop wise added to the cells that were further incubated at 28°C for 16 to 24 hours. Cells were further washed twice by centrifugation for 10 minutes at 100 x g followed by resuspension in complete medium. Fresh complete medium was added to the medium for incubation for 48 to 72 hours before cell collection.

2.4.1.c- Stable transfection of SC2 cells

Transfections were carried out in the conditions described before by adding 19µl of the expression vector recombinant for the target gene and 1 µg of the pCoBlast vector for selection of the transfected cells to the solution A. After 16 to 24 hours incubation of the cells with the calcium phosphate precipitate containing DNA, cells were washed as described for transient transfections. Selection of transfected cells was carried out by adding 25µg/ml of blasticidin to the complete medium. Clones were observed after 3 weeks of selection.

2.4.2- Mammalian cells: porcine PK15 cells

PK15 cells are adherent pig epithelial renal cells that grow in MEM supplemented by 10 % fetal calf serum, 10mM Hepes buffer, 2mM glutamine and streptomycin/penicillin antibiotics. Transfections were carried out using the Lipofectamine™ 2000 (Invitrogen, USA) with 2µg of expression vector for a 35 mM dish containing cells grown at 70% confluence. Transfection conditions were as recommended by the manufacturer. Cells were collected after 48h of transfection.

2.5- RNA extraction from transfected cells

Total RNA was extracted from SC2 transfected cells using the RNeasy Mini Kit (Qiagen, USA). RNA concentrations were determined by Nanodrop quantification (Thermo Fisher Scientific, USA) and the quality was assessed by the Agilent 2100 Bioanalyzer (Agilent Technologies, Germany).

2.6- Monoclonal antibody production specific for the SLA-Ib molecules

Antibodies were produced by the P.A.R.I.S. Anticorps company (Compiègne, France) that developed a peptide and a genic strategy for immunization of mice. We provided the sequences of SLA-6, -7 and -8 available in the supplementary files S1 and S2. All reports dealing with the production of antibodies are in supplementary files S3 and S4.

2.7- Western blot analysis for detection of expressed proteins in SC2 cells

SC2 cells (5×10^6) were collected by centrifugation at 1,500 g for 5min at room temperature, washed with PBS and centrifuged again. Cell pellets were incubated in 400 μ l of lysis buffer (5 mM TrisHCl pH 7.4, EDTA 2mM, Triton 1%, PMSF 0.1% and protease inhibitors) for 1h on ice. Cell lysates were further centrifuged for 40 min à 4°C and the supernatants were collected and transferred in a new tube. The proteins were separated by SDS-PAGE Gel Electrophoresis (12% acrylamid) and the gels were blotted onto a nitrocellulose membrane (Hybond, Amersham). The membrane was incubated for 30 min in 20 ml of a solution of

Tris-EDTA-NaCl buffer (TBS: 25 mM Tris-HCl, 1 mM EDTA, 150 mM NaCl, pH 7.6) containing 5% bovine serum albumin (BSA). The membrane was then incubated for 30 min in 20 ml of a solution of Tween-TBS (TBS with 0.2% Tween-20) containing 2.5% BSA and the primary antibody at the appropriate dilution. The membrane was further washed three times for 10 min in Tween-TBS before incubation for 30 min at room temperature in 20 ml of a solution of Tween-TBS containing 2.5% BSA and a horse radish peroxidase-conjugated secondary antibody at the appropriate dilution. The membrane was then washed three times for 10 min in Tween-TBS followed by incubation for 6 min in a solution containing 2 ml peroxydase and 2 ml Luminol/Enhancer solution for further detection of the target proteins conjugated with specific antibodies by chemiluminescence. The membrane was covered by a plastic film and placed in a cassette with an autoradiography film.

2.8- Fluorescent Activating Cell Sorting (FACS) analysis

2.8.1- Detection of molecules on the cell surface

Since SC2 cells are semi adherent cells, no special protocol was developed to detach them from the plastic surface of the culture flasks. Conversely, the adherent PK15 cells were isolated by incubation with 5 mM EDTA instead of Trypsin/EDTA in order to preserve the integrity of membrane-bound molecules.

Cells were collected and washed in 50 mL D-PBS without $MgCl_2$ and $CaCl_2$, incubated with 2 ml pig serum at 4°C for 20 min, washed in 50 mL D-PBS without $MgCl_2$ and $CaCl_2$ and then in 50 mL S/W buffer (1 g/L NaN_3 , 10 g/L bovine serum albumin in PBS, pH 7.3). The cell pellets were resuspended in S/W buffer at a final concentration of 5×10^6 cells/ml.

10⁶ cells were used for each antibody labelling. The antibodies that were used are presented in Table VIII. Cells were stained with primary mAbs at the appropriate dilution for 25 min at 4°C, washed in S/W buffer and stained with a conjugated secondary antibody at the appropriate dilution for 25 min at 4°C. The combinations between primary and secondary antibodies are presented in Table VIII. After washing in S/W buffer, cells were fixed in BD Cellfix solution (Becton Dickinson, Germany). Data acquisition and analysis were carried out with the FACScan and CELLQuest software (Becton Dickinson, UK).

Table VIII: Antibodies used for FACS analysis

Primary antibodies				Secondary antibodies	
Name	Specificity	Isotype	Provider	Specificity	Provider
87G	HLA-G	IgG2a, mouse	INRA, VIM	Phycoerythrin-conjugated anti-mouse IgG2a	Southern Biotech, France
74-11-10	SLA Ia	IgG2b, mouse	VMRD, France	FITC-conjugated anti-mouse IgG2b	Southern Biotech, France
B2M-02	B2M, cell surface	IgG2a, mouse	Abcam, France	Phycoerythrin-conjugated anti-mouse IgG2a	Southern Biotech, France
2M2	B2M	IgG1, mouse	Abcam, France	Allophycocyanin-conjugated anti-mouse IgG1	BD Biosciences, France
in test*	SLA-6 or SLA-8	not determined, mouse	P.A.R.I.S., France	FITC-conjugated anti mouse IgG (GAM, 115-096-006)	Interchim

* Monoclonal antibodies provided by the company P.A.R.I.S.: 24 and 23 clones to screen for SLA-6 and -8, respectively.

2.8.2- Detection of molecules in the cytoplasm by cell permeabilization

We used the KIT Fixation and Permeabilization CALTAG (Invitrogen, France, Reference GAS-004) and followed the recommendations of the manufacturer. This protocol was used to screen the monoclonal antibodies provided by P.A.R.I.S.

3- Polymorphism studies

A scheme summarizing the strategy developed to characterize the SLA Ib polymorphism as presented in Figure XIV. Long range PCR and sequencing were the major methods to identify SNPs, and pyrosequencing was the chosen method to validate SNPs and confirm hypotheses on gene duplication.

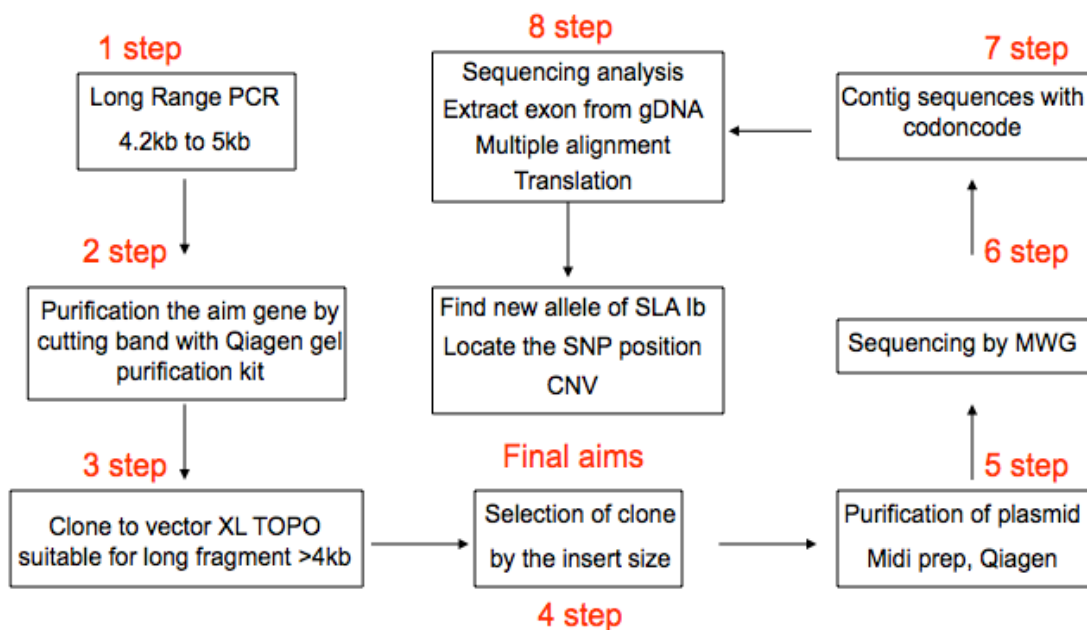


Figure XIV: General strategy to study the polymorphism of SLA-Ib genes

3.1- Primer sets for long range PCRs specific for the SLA-Ib genes

Primer pairs suitable to amplify the SLA-6, -7 and -8 genes from upstream to the promoter to downstream to the polyadenylation site were designed using the Primer3 program available online at <http://frodo.wi.mit.edu/primer3/> (Table IX). The primers were designed from the BAC clone that comprises the SLA-6, -7 and -8 genes (AJ251914.1) (Chardon et al. 2001).

Table IX: Primer sets designed to amplify the SLA Ib genomic sequences

Gene	Primers	Sequences (5' to 3')	Size (bp)
SLA6	SLA6_pr_F	TCAGGGCTATGGAATGAAGG	5208
SLA6	SLA6_pr_R	ACCATAGCGGGAACCTCCTCT	
SLA7	SLA7_pr_F	ATGCCTTCTACCCTCCTGGT	4145
SLA7	SLA7_3UTR_R2	ACTGGATCAGCAGTGGCTCT	
SLA8	SLA8_pr_F	GCCTCCATAATGATCGCTGT	4436
SLA8	SLA8_pr_R	ACTCAGCACAGCAAATGTGG	

3.2- Animals

Genomic DNA of eight MeLiM pigs corresponding to either founders (B52, C284, C321, F206, and F213) or progeny (484, 485, 486) of MeLiM pigs and one Large White pig with the Hp1a.0 haplotype were included in the study.

3.3- Long Range PCR and cloning

For PCR, the long range PCR kit was used (Qiagen, USA) that combines a powerful polymerase blend with an innovative buffer system designed for efficient amplification of long targets up to 40 kb. The reactions were set up as recommended using 100 ng template genomic DNA and the following thermocycling conditions were used: initial denaturation for 3 min at 93°C, (93°C 30s; 62°C 1 min; 68°C 5 min) x 35 cycles. The PCR products were cloned using the TOPO XL PCR cloning kit (Invitrogen, USA), which provides a highly efficient, 5 minute one-step cloning strategy for the cloning of long PCR products generated by commercial enzyme mixes specifically formulated to generate long PCR products. Ligated products were clone in DH10B bacteria as previously reported in section 1.7.3.

3.4- DNA preparation and sequencing

In order to sequence the full sequence of the cloned inserts on both DNA strands, a set of 12, 11 and 14 primers were designed for SLA-6, -7 and -8 genes respectively (Table X). For this sequencing purpose, large amounts of plasmid DNA were prepared using the Midi Plasmid Purification kit by following the recommendations of the manufacturer (QIAGEN, USA). Sanger sequencing was performed by the MWG company.

3.5- Sequence analysis

Sequence edition and contig construction were done using the CondonCode software Version 3.5 (<http://www.codoncode.com/>). The sequences of each allele were assembled by overlapping forward and reverse sequence fragments. All sequences were compared to the reference sequence provided by the haplotype Hp1a.0 (Renard et al. 2006).

Table X: Primers for SLA Ib gene sequencing

Gene	Primer	Sequence (5' to 3')	Position
SLA-6	SLA6_pr_F	TCAGGGCTATGGAATGAAGG	5'UTR
	SLA6_e1_F1	GGAAGGATGCAGGTCACG	exon1
	SLA6_in2_F1	CTCACCTCCTTCTCCCTTCC	intron2
	SLA6_e2_F	GGCCACGGTAGTGACCTTTA	exon2
	SLA6_e4_F	CTTCTGGAGAGGAGCAGAGC	exon4
	SLA6-exon5_F	GTGCTTGTGGTCACTGTGCT	exon5
	SLA6_e6_F	TGGAAATAGAGGGAACCTATGTTTCAG	exon6
	SLA6_exon2_R1	AAGTTCCTGCACCCCCTTAC	exon2
	SLA6_e4_R	ACTCTAAGATGCTGGGCCCT	exon4
	SLA6_in3_R	AAGACCACTGGAACCTGTGG	intron3

Gene	Primer	Sequence (5' to 3')	Position
SLA-7	SLA6_pr_R	ACCATAGCGGGAACCTCCTCT	3'UTR
	SLA6_pr_R2	AGCAATCACTGTACAGCAAAGTG	3'UTR
	SLA7_pr_F	ATGCCTTCTACCCTCCTGGT	5'UTR
	SLA7_pr_F2	AGACAAACTCAGGGCATGGA	5'UTR
	SLA7_e1_F1	ATGGGGCCCCGAGCCCTCCT	exon1
	SLA7_i3_F1	GACGGAGGCATCTATCCAAA	intron3
	SLA7_e4_F	TGGAGAGGAGCAGAGCTACA	exon4
	SLA7_i5_F	TCTTGCCTTGGGTCTGAGAT	intron5
	SLA7_in6-F	GCTGAGATCCCCAAAACCTT	intron6
	SLA7_exon2_R1	ATCCAGGTCACGACCATCCT	exon2
	SLA7_3UTR_R	GCAGCTGTTTCCCCTATAG	3'UTR
	SLA7_3UTR_R2	ACTGGATCAGCAGTGGCTCT	3'UTR
	SLA7_pr_R	TCTCATTCCTGGCAATACC	3'UTR
SLA-8	SLA8_pr_F	GCCTCCATAATGATCGCTGT	5'UTR
	SLA8_pr_F2	CCCAAGAAATGGCAAAAAGA	5'UTR
	SLA8_e1_R1	ATGGAGTCTCAGATGCTTCTTC	exon1
	SLA8_exon2_F1	GAGGGAAGGGTCTCAACCTC	exon2
	SLA8_i3_F	GTGACTAGCCCCAAAGGACCA	intron3
	SLA8_e4_F	CCTGGAGAGGAGCAGAGCTA	exon4
	SLA8_i5_F	GCCTCTCACAGGGTGTTTTTC	intron5
	SLA8_i6_F	GATGGGGAATGTCAGGGAAT	intron6
	SLA8_R	AGAAGCCAATCAAGTCACCG	3'UTR
	SLA8_exon2_R1	TAGTCTCCCCCATCTCCCAC	exon2
	SLA8_i3_R	GTCAGGAGAGGAGGGATTCC	intron3
	SLA8_exon4_R	TCACCCTGAGATGGGGTAAG	exon4
	SLA8_i6_R	CAAGTAAGTAGGGGAGGGGG	intron6
	SLA8_pr_R	ACTCAGCACAGCAAATGTGG	3'UTR

3.6- Pyrosequencing

Pyrosequencing-based genotyping require three primers that comprise two PCR primers (one primer is biotin-labeled for immobilization to sepharose beads), and one sequencing primer, which be anchored close to polymorphic position. All primers were designed with the PSQ design software (Biotage, <http://www.biotage.com/>) and were selected to be 18 - 24 bases long with a T_m of 62 - 68°C. Three pairs of SLA-7 specific PCR primers were designed that targeted four polymorphic positions (Table XI).

Table XI: Primers for pyrosequencing PCR and sequencing

Position	Primers	Sequences (5' to 3')	Size
Forward	SLA7_1535_F	TGCTCCGGTTCTCTTCGAC	56bp
Reverse	SLA7_1535_RB	GGCCCCCAAGAGATCTAT	
Primer for sequencing	SLA7_1535_S	TCTCTTCGACCCTGAC	
Forward	SLA7_3601_F	GGAGCAGCCGTAGCGAGAA	117bp
Reverse	SLA7_3601_RB	ACCCCCATAACCCTGTTGTGTC	
Primer for sequencing	SLA7_3601_S	CGAGAACTCTGGTGTATGT	
Forward	SLA7_3974_F	CGTGAGGATGCAAGTTTGATC	79bp
Reverse	SLA7_3974_RB	TGGCCTACACCACAGCTCA	
Primer for sequencing	SLA7_3974_S	GATCCCTGGCCTCAC	

The first pair of primers (SLA7_1534_F and biotinylated SLA7_1355_RB) targeted the positions 1509 to 1564 (56 bp) within SLA-7 intron 2. Two SNPs (A1535G and G1540T) were identified in this segment. Genotyping was conducted using sequencing primer (SLA7_1535_S) and dispensation order of nucleotides was determined using PSQ software (Biotage). The second pair of primers (SLA7_3601_F, SLA7_3601_RB) target 117 bp of SLA7 exon 7, in order to genotype the SNP at position 3601 (C < T), with the sequencing primer SLA7_3601_S. The third pair of primers (SLA7_3974_F, SLA7_3974_RB) matched

SLA7 3'UTR, to genotype an insertion/deletion polymorphism located from 3974 to 3979, using sequencing primer SLA7_3974_S. The PCR product might be 73 bp or 79 bp according to allele.

Genomic DNA samples were amplified with the PCR primer pairs previously designed to amplify segments in which SNPs have been identified. All PCR products were checked by agarose gel electrophoresis before pyrosequencing. The PCR products were denatured to single-stranded templates, captured with sepharose beads coated with Streptavidine, washed and transferred in sequencing reaction tubes containing sequencing buffer and the sequencing primer. Pyrosequencing reactions were performed with the PyroMark Q24 system, by sequential injection of nucleotides according to dispensation order. Results were analyzed with the Biotage software. The relative frequency of the paralogous sequences is calculated using the Pyrosequencing AQ software.

The pyrosequencing protocol has been described below by highlighting the various constraints there were considered in each step.

For the primer design: three primers are used for pyrosequencing. A primer pair is used to amplify the target PCR product. The third primer is used to sequence the PCR product. The primer in opposite orientation by comparison to the sequence primer should be biotin labelled for immobilisation to sepharose beads and this biotinylated primer should not include hairpin loops or duplexes, which could cause background in pyrosequencing assay. For SNP analysis, the sequencing primer should be positioned within 5 bases of the SNP. Ideally, there would be one base

between the primers and SNP of interest, but the shorter the read, then the quicker the run time.

PCR amplification: See Table XII for the PCR reaction. After incubating at 95 °C for 5 min, the amplification reaction was carried out for 50 cycles with the following cycle: denaturation step for 30 s at 95°C, annealing step for 30 s at 60°C, and extension at 72°C for 10s.

Table XII: The mixture for PCR amplification

Reagent	Volume (μl)	Final concentration
Genomic DNA (40 ng/μl)	1	40 ng
10× PCR buffer	2.5	1× (contains 1.5 mM MgCl ₂)
MgCl ₂ (25 mM)	1.6	1.6 mM
dNTP mix (8 mM)	1.25	200 μM
Primer forward (10 pmol μl ⁻¹)	0.5	0.2 μM
Primer reverse (10 pmol μl ⁻¹)	0.5	0.2 μM
Taq DNA polymerase (5 U/μl)	0.4	2 U
Water	—	up to 25 μl

Pyrosequencing Preparation: 18 μl of PCR product was used per pyrosequencing reaction as recommended by the manufacturer (PyroMark).

Results

As presented in the introduction section, MHC Class Ib genes are usually expressed in a tissue specific manner and may present spliced variants, as mainly reported in human. Moreover, their polymorphisms are very limited compared to MHC class Ia genes. When we started the work in swine, it had been reported a less tissue-restricted transcription of SLA-Ib genes compared to HLA-Ib genes (Crew et al. 2004) and the following data were known on gene and RNA structures: i)- SLA-6, -7 and -8 genomic sequences from the Hp1a.0 haplotype (Chardon et al. 2001, Renard et al. 2006, VEAG_database), ii)- RefSeq RNAs corresponding to putative full length coding sequences (Crew et al. 2004, VEGA_database), iii)- one alternative variant for SLA-6 with a missing exon 3 (VEAG_database). For protein expression and functional studies, no antibody specific for any SLA-Ib proteins had been described. In order to increase knowledge on SLA-Ib transcription, our aim was therefore to address three complementary questions related to RNA transcription, protein expression and gene polymorphism.

1- SLA Ib transcription

As presented earlier in this manuscript, we carried out experiments in order to characterize SLA-Ib transcription patterns and to extend studies on tissue specificity. The addressed questions could be summarized as follows: i) - do SLA-6, -7 and -8 express alternative variants and if yes, which ones? ii) - in which tissues are the RNAs expressed?

A publication on SLA-7 transcripts is in press (Publication #1). We have prepared a manuscript reporting the whole set of variants for the three SLA-Ib genes. However, since monoclonal antibodies that recognize SLA-6 and SLA-8 have been recently made available and in test in the frame of my PhD, it has been decided to delay the result submission in order to include results on protein expression at the cell surface if the antibody screening is successful. Results on SLA-Ib tissue specificity transcription are in press in Animal Genetics (Publication #2). The two publications are at the end of the Results section together with abstracts and posters presented in several national and international meetings during my project (three abstracts and three posters).

1.1- Characterization of SLA-Ib transcripts

1.1.1- Identification of alternative transcription for SLA-6 and -7 but not SLA-8

The transcripts were characterized by RT-PCR as described in the Materials and Methods section using heart, thymus and brain RNAs from MeLiM animals. Primers suitable for amplification from exon 1 to the 3'UTR clearly showed that several DNA fragments were amplified for SLA-6 and -7 whereas a unique band was amplified for SLA-8 (Figure XV).

Additional primers suitable for amplification from exon 4 to the 3'UTR confirmed several bands for SLA-6 and -7 and a unique band for SLA-8 (see Figures XIX, XX and XXI)

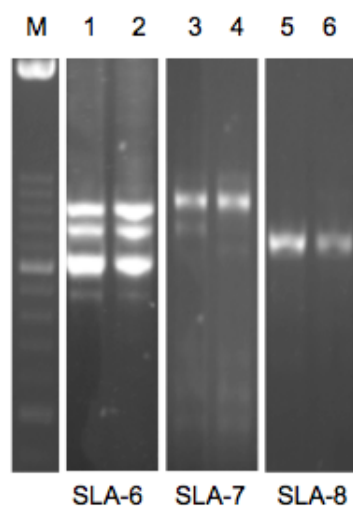


Figure XV: PCR result of full-length cDNA of SLA Ib

RT-PCRs from thymus (lanes 1, 3 and 5) and heart (lanes 2, 4 and 6) with primers suitable for amplification from exon 1 to the 3'UTR that is specific for SLA-6 (1,2), SLA-7 (lanes 3 and 4) and SLA-8 (lanes 5 and 6). M: DNA ladder.

1.1.2- SLA-6: five transcripts

Five SLA-6 transcripts were identified in MeLiM pigs and were further called SLA-6, SLA-6-1, SLA-6-2, SLA-6-3 and SLA-6-4 (Figure XVI). The five sequences have been submitted to EMBL/DDBJ/NCBI and were given the accession numbers GU322911 to GU322915₂. Respectively, among the five transcripts, SLA-6 and SLA-6-1 had been previously reported and the three transcripts SLA-6-2, SLA-6-3 and SLA-6-4 stand for newly identified spliced variants.

1.1.2.a- SLA-6 transcript: a transcript encoding a protein of 7 exons

The SLA-6 transcript is 1221 nt long and comprises 7 exons that encode a protein of 370 amino acids (AA). This transcript is similar to the transcript referred to as SLA-6-001 in the VEGA database and stand for the full coding sequence.

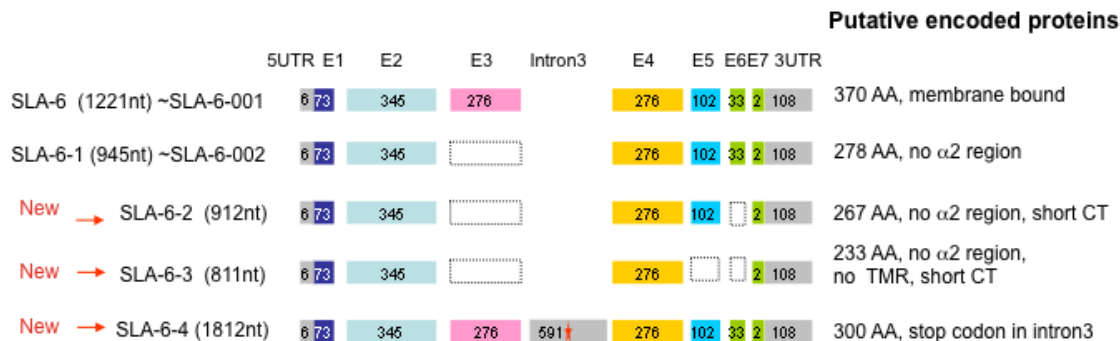


Figure XVI: Five SLA-6 transcripts obtained from MeLiM pig

The transcript names with their size in nucleotides (nt) are on the left hand of the scheme. Exons and introns are represented by boxes with colours that specify the molecule functional domains (dark blue: leader peptide, pale blue: alpha 1 domain, pink: alpha 2 domain, yellow: alpha 3 domain, bright blue: transmembrane domain, green: cytoplasmic tail; grey: non coding sequences). The sizes are indicated in nt within each box. Missing exons are represented by empty boxes. The red star in the intron3 of SLA-6-4 stands for a stop codon. Size and features of the putative encoded proteins are summarized on the right hand of the figure.

1.1.2.b- SLA-6-1 to SLA-6-4 transcripts: alternative spliced variants

SLA-6-1

The SLA-6-1 transcript is 945 nt long and its structure is similar to the transcript referred to as SLA-6-002 in the VEGA database (Transcript ID: OTTSUST000000000780). The striking feature of this transcript is the absence of exon 3 (Figure XVI), meaning that the putative encoded protein would be 278 AA long and would lack the alpha 2 domain.

SLA-6-2

The SLA-6-2 transcript is 912 nt. Exons 3 and 6 are spliced, leading to a protein of 267 AA. The putative protein would lack the alpha 2 domain and harbour a cytoplasm tail with a single AA encoded by exon 7 (see Figure XVI).

SLA-6-3

Among the five SLA-6 transcripts found in this study, the SLA-6-3 transcript is the shortest with 811 nt. Exons 3, 5 and 6 are spliced. The putative encoded protein comprises 233 AA and lacks the alpha 2 domain, the transmembrane domain and harbours a cytoplasmic tail with a single AA as the transcript SLA-6-2 (see Figure XVI).

SLA-6-4

Among the five SLA-6 transcripts found in this study, the SLA-6-4 transcript is the longest with 1812 bp. This transcript includes the seven exons and intron 3 (591 nt) is retained. The putative encoded molecule comprises 300 AA that only code for alpha 1 and 2 domains due to a premature stop codon within intron 3 (see Figure XVI).

1.1.3- SLA-7: seven transcripts

A total of seven transcripts are now characterized for the SLA-7 gene by including the available full length RefSeq RNA and the six RNA variants that were sequenced from brain and thymus tissues in our study (RNAs referred to as SLA-7 and SLA-7-1 to SLA-7-5) (Figure XVII). Surprisingly, all identified transcripts were new, including the transcript corresponding to the full coding sequence. SLA-7 stands for the full coding sequence (1465

nt, ACC=GU322918), and the five others for RNA variants characterized either from exon 1 to 3'UTR such as SLA-7-1 (1366 nt ACC=HQ224544), SLA-7-2 (1443 nt, ACC=HQ224547), SLA-7-3 (1101 nt, ACC=HQ224546), SLA-7-4 (1443 nt, ACC=HQ224545) or from exon 4 to UTR such as SLA-7-5 (464 nt, ACC=GU322919). Parts of these results are reported in publication 1.

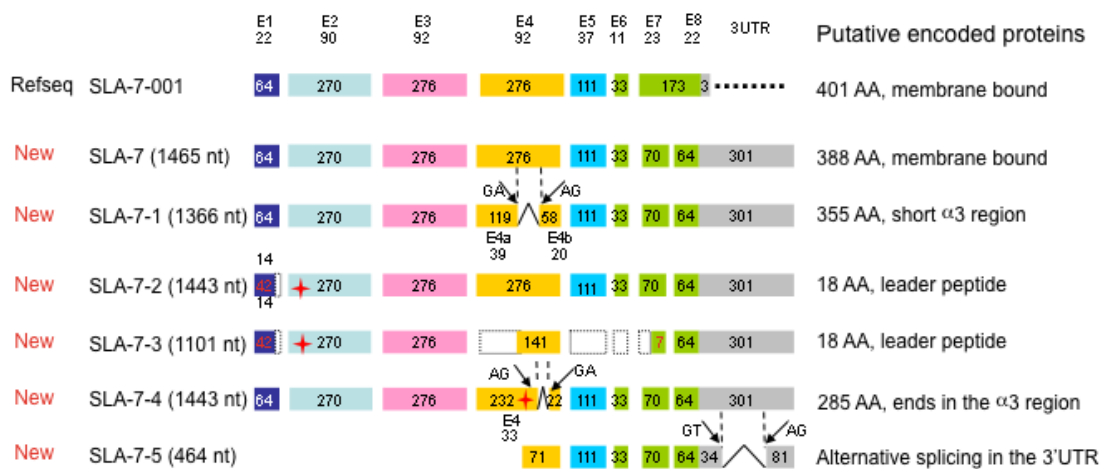


Figure XVII: SLA-7 transcripts obtained from MeLiM pig

The transcript names with their size in nt are on the left hand of the scheme. Exons and introns are represented by boxes with colours that specify the molecule functional domains (dark blue: leader peptide, pale blue: alpha 1 domain, pink: alpha 2 domain, yellow: alpha 3 domain, bright blue: transmembrane domain, green: cytoplasmic tail; grey: non coding sequences). The sizes are indicated in nt within each box. Missing exons are represented by empty boxes. Red stars represent premature stop codons. The di-nucleotides present in the genomic sequence at the splice sites are indicated for the variants for transcripts SLA-7-1, SLA-7-4 and SLA-7-464. Size and features of the putative encoded proteins are summarized on the right hand of the figure.

1.1.3.a- SLA-7-001 and SLA-7: two different transcripts that stand for the full coding sequence

The reference RNA for the SLA-7 full coding sequence is referred to as SLA-7-001 and encodes seven exons that lead to a protein of 401 AA (Figure XVII). By contrast, we identified a transcript that comprises eight exons and encodes a shorter protein of 388 AA. Both putative full length molecules differ in the cytoplasmic tail that is shorter in the isoform encoded by the SLA-7 RNA we have identified. Strikingly, the exon 8 identified in RNA SLA-7 is 64 nt long instead of 1 or 2 as found for the last exon of most class I transcripts. The cytoplasmic tail of SLA-7 molecules is longer than for other class I molecules. We have to mention that we could never find the previously known transcript SLA-7-001 that stands for the SLA-7 RefSeq sequence.

1.1.3.b- SLA-7-1 to SLA-7-5: alternative RNA variants with a complex pattern

SLA-7-1: a variant spliced within exon 4

The transcript SLA-7-1 comprises nine exons due to a splicing event that occurred within exon 4, leading to exons that were further named exons 4a and 4b (Figure XVII). The SLA-7-1 encodes a protein of 355 AA that harbours a shorter alpha 3 domain than the full length isoform with 59 AA instead of 92 AA. The splicing junctions were much unexpected with rarely used donor/acceptor splice sites GA-AG instead of the canonical ones GT-AG. It has been shown that the GA-AG splicing site is rarely used and a few cases have been reported among which splicing in the human parafibromin gene (Bradley et al. 2005). Our results

suggest that the SLA-7 gene may be subject to subtle regulation resulting in the use of rarely used non canonical splicing sites.

SLA-7-2 and SLA-7-3: two variants with a shorter exon 1

Both SLA-7-2 and SLA-7-3 RNAs harbour a splicing event between exons 1 and 2 that leads to a short exon 1 of 42 instead of 64 nucleotides and creates a premature stop codon in exon 2 (Figure XVII). The putative encoded molecule would stand for a leader peptide of 18 AA. The two RNAs differ in the part that spans from exon 4 to exon 7. The variant SLA-7-2 is similar to the full length SLA-7 RNA with no additional splice downstream to the exon 1. Conversely, the RNA SLA-7-3 presents several splicing events that result in a transcript with shortened exons 4 and 7 and missing exons 5 and 6.

SLA-7-4: a variant that could be related to a deletion in genomic DNA

This variant presents a splicing event within exon 4 that creates a premature stop codon in exon 4. The putative encoded molecule would be 285 AA long and would comprise full alpha 1 and 2 domains and a short alpha 3 domain (Figure XVII). The donor/acceptor splice sites (AG-GA) are not canonical. In contrast to the RNA variant SLA-7-2, we could not find any data reporting the use of this non canonical spliced site. We are not convinced that this RNA is a true sliced variant. Indeed, it may result from a deletion in genomic DNA. We have not further studied this RNA but our ongoing experiments on SLA-7 polymorphism suggest copy number variations according to haplotypes. The transcript SLA-7-4 may be expressed from an SLA-7 copy harbouring a deletion that could explain the putative splice site reported here. We have however no conclusion on this hypothesis yet.

SLA-7-5: a spliced variant in the 3'UTR

The SLA-7-5 was identified with primers suitable for amplification from exon 4 to the 3UTR. A splicing event was identified in the 3UTR, 31 nt downstream to the termination codon. The encoded molecules are not modified by the splicing event. The canonical GT-AG rule was used for this splicing (Figure XVII).

1.1.4- SLA-8: a unique full length transcript

In our study, a single transcript (SLA-8, GU322909) was found with a structure similar to the reference sequence SLA-8-001 (TranscriptID: OTTSUST00000000786). The RNA stands for a full coding sequence with eight exons that encode a protein of 355 AA (Figure XVIII).

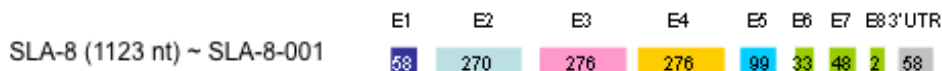


Figure XVIII: SLA-8 transcript obtained from MeLiM pig

The SLA-8 transcript is 1123 nt long and has a structure similar to the reference SLA-8-001 transcript (VEGA database).

1.1.5- Validation of two splice sites by genomic DNA amplification of SLA 1b

We have identified many RNA variants for SLA-6 and SLA-7. In order to check that their expression is due to true splicing events and not pre-existing deletions in genomic DNA, we have designed primers suitable for amplification of genomic DNA. For SLA-6, we targeted

exons 2 to 4 in order to validate the splicing of exon 3 in transcripts SLA-6-1, -2 and -3 (Figure XVI). For SLA-7, we targeted the genomic segment from intron 6 to the 3UTR to validate the splicing event found in the 3UTR. In all cases, a unique genomic band was found (Figure XIX), demonstrating that the variants SLA-6-1, -2, -3 and SLA-7-5 are due to splicing events and not to deletion in the genomic DNA. As mentioned before, we have not designed primers targeting the possible splicing event found in exon 4 for the variant SLA-7-4 and the splicing event is not validated.

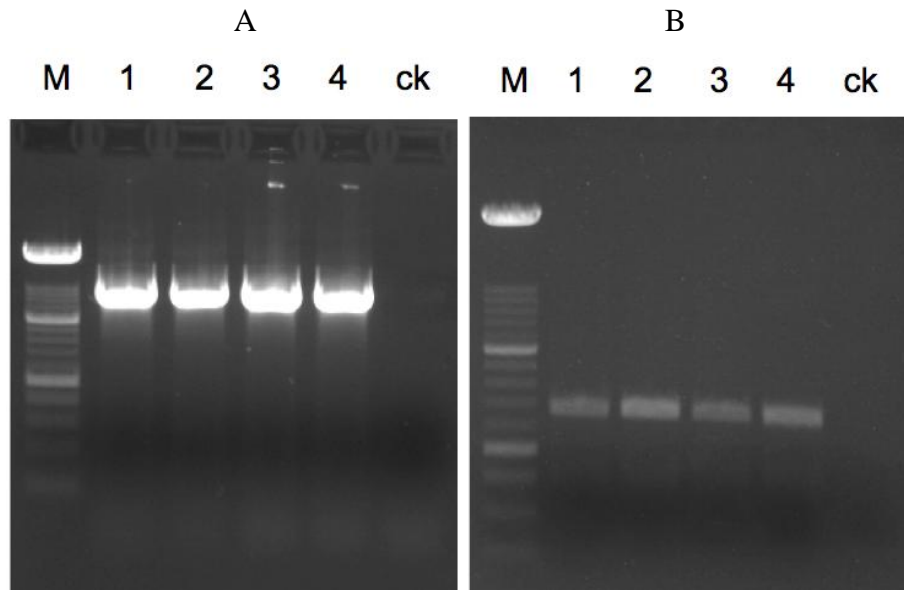


Figure XIX: Validation of splicing events by PCR amplification of SLA-6 (A) and SLA-7 (B) partial segments from genomic DNA

M: DNA ladder (XIV), 1: MeLiM (#484), 2: MeLiM (#485), 3: MeLiM (#486), 4: Large White, ck: negative control. For SLA-6, the primers amplify a fragment from exons 2 to 4. For SLA-7, the primers amplify a fragment from intron 6 to the 3UTR.

1.2- Tissue specificity

The work on SLA-Ib expression in a wide range of tissues had started before my PhD project and initially included analysis of the transcription levels by targeting a junction between exons 2 and 3 without knowing that splicing events could occur and by comparing the relative levels of SLA-Ib transcripts to those of SLA-Ia transcripts (Publication # 2). In a second part of the work, I studied the expression of various splice variants by RT-PCR in the same range of tissues (results partially presented in Publication #1).

1.2.1- Relative expression of SLA-Ib genes and comparison with SLA-Ia genes

Twenty-five adult tissues from MeLiM and Large White pigs, pig renal PK15 cells infected with the Pseudorabies virus and peripheral blood mononuclear cells (PBMCs) stimulated by lipopolysaccharide or a mixture of phorbol myristate acetate and ionomycin were included in our study. Relative transcription was quantified by quantitative real-time PCR. On average, in adult tissues and PBMCs and compared to SLA-6, the transcription level of SLA-Ia genes was 100–1000 times higher, the level of SLA-8 was 10–20 times higher, and that of SLA-7 was five times higher. Thus, SLA-8 is the most transcribed SLA-Ib gene, followed by the SLA-7 and SLA-6 genes. The highest transcription levels of SLA-Ib transcripts were found in the lymphoid organs, followed by the lung and the digestive tract. The tissue variability of expression levels was widest for the SLA-6 gene, with a 1:32 ratio between the lowest and highest levels in contrast to a 1:12 ratio for the SLA-7 and SLA-8 genes and a 1:16 ratio for the SLA-Ia genes. During PK-15 infection and PBMC stimulation, SLA-Ia and SLA-8 genes were down-regulated, whereas SLA-6 and SLA-7 were up-regulated, down-regulated or not

significantly modified. Our overall results confirm the tissue-wide transcription of the three SLA-Ib genes and suggest that they have complementary roles (see Publication 2).

1.2.2- Expression of the transcript variants in various tissues

Since RNA variants were found for SLA-6 and SLA-7 but not SLA-8, we were interested in studying their presence in various tissues of Large White and MeLiM pigs by RT-PCR. We used the same set of tissues as the set that had been created to study the relative transcription of SLA-Ib genes by qRT-PCR. RT-PCRs were carried out in order to amplify cDNAs from exon 4 to the 3UTR.

1.2.2.a- SLA-6

Three different PCR products are expected by amplifying cDNAs with the primers SLA-6-e4-F and SLA-6-3UTR-R (Table VI). Amplicons of 320 bp are expected for the variants SLA-6, SLA-6-1 and SLA-6-4, amplicons of 287 bp are expected for the variant SLA-6-2 and amplicons of 187 pb are expected for the variant SLA-6-3 (see Figure XX).

We detected a band of 320 bp in all tested tissues (Figure XX) and this band had the strongest intensity. This finding indicates that SLA-6 transcripts harbouring a full cytoplasmic tail are predominantly expressed in all tested tissues (SLA-6, SLA-6-1, SLA-6-4). In our experimental design we could not make the difference between SLA-6 transcripts lacking or not exon 3 and cannot directly conclude on a predominant expression of SLA-6 full coding sequence transcripts.

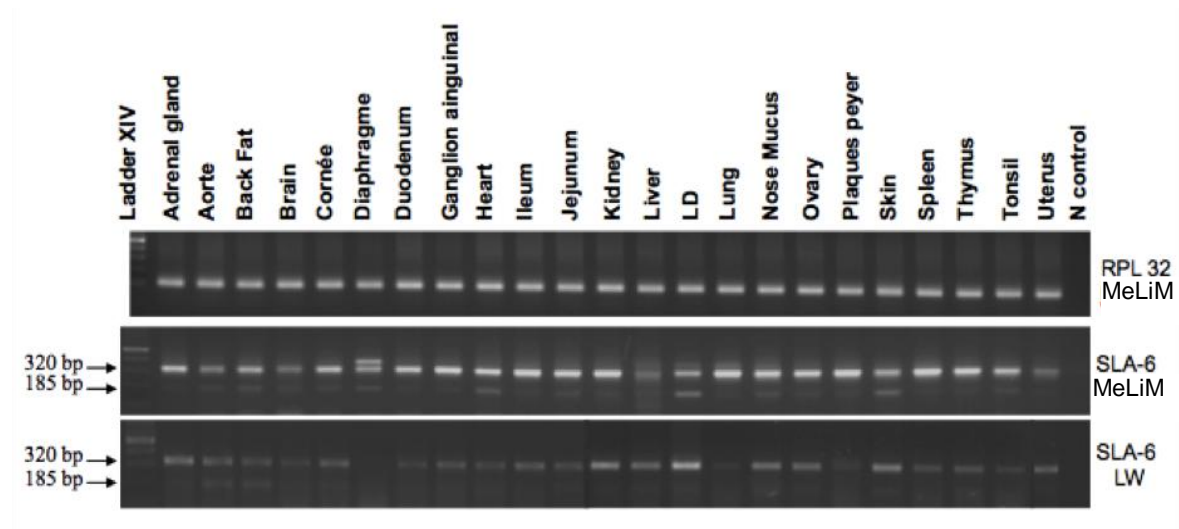


Figure XX: Detection of SLA-6 RNA variants in various tissues from MeLiM and Large White pigs by RT-PCR

PCR primers were chosen to amplify cDNAs from exon 4 to the 3UTR. The RPL32 was chosen as an internal control for transcription levels. Tissue types are indicated on top of the figure. Sizes of the amplicons are indicated for bands that are specific to SLA-6 RNAs.

A band of 185 bp was also found in most tissues but with various intensities and ratios between this band and the band of 320 bp was not the same in all tissues. The amplicons of 185 bp target the SLA-6-3 transcript that harbours a shortened cytoplasmic tail (no exon 6) and no transmembrane domain (no exon 5). Absence of the transmembrane domain suggests that the encoded molecule may be soluble. The SLA-6-3 transcript seems to be slightly expressed in many tissues with a higher expression in heart, longissimus dorsi (LD) muscle and skin as shown in the MeLiM pig (Figure XX).

The band of 287 pb was hardly detectable whatever the tested tissue, indicating that the variant SLA-6-3 is the less expressed.

More experiments are required to generalise these results but we already show that some SLA-6 variants are more widely expressed than others and that relative expression of SLA-6 RNA variants within tissues differ according to tissues.

1.2.2.b- SLA-7

PCR primers were chosen for amplification of cDNAs from exon 4 to the 3UTR (primers SLA-7-e4_F/SLA-7-3UTR_F, Table VI). Two main PCR products of 464 and 650 bp were found (Figure XXI) that correspond to amplification of the subset of variants that are spliced (SLA-7-5) or not (SLA-7, SLA-7-2, SLA-7-4) in the 3UTR (Figure XXI). The primers were not suitable for amplification of cDNAs spliced in exon 4.

In MeLiM pigs, the band intensity between the amplicons 464 bp and 650 bp varied according to tissues. The 650 bp long amplicon was predominant in most tissues. However, the ratio was approximately 1:1 in cornea, inguinal lymph node, lung, ovary, Peyer's patches, thymus and tonsil. Only the 650 bp long amplicon was found in the diaphragm and only the 464 bp long band was detected in the longissimus dorsi muscle.

In MeLiM pigs, two bands of 464 and 650 bp were found in almost all tested tissues whereas in Large White tissues, only amplicons of 650 bp could be found. If variants spliced in the 3UTR are expressed in Large White pigs, the expression level is much lower than in MeLiM tissues.

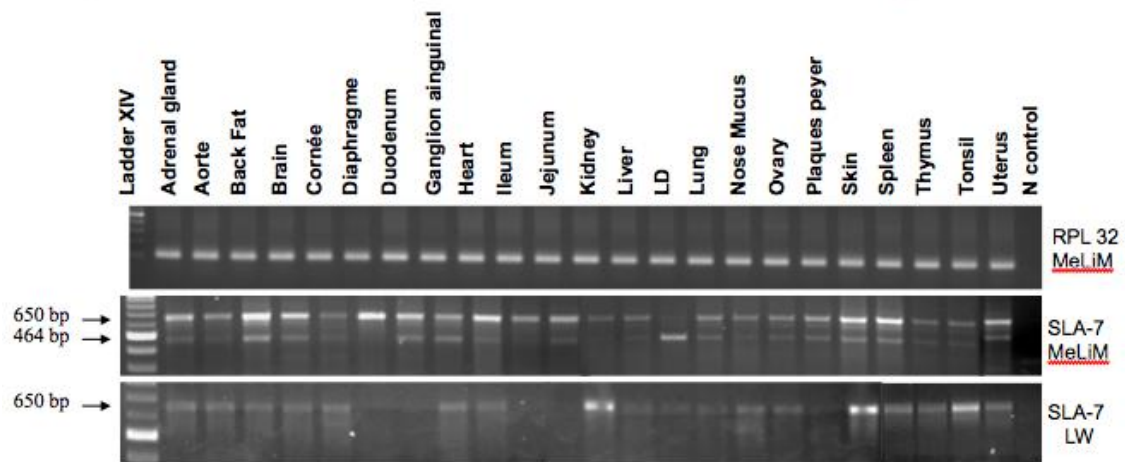


Figure XXI: Detection of SLA-7 RNA variants in various tissues from MeLiM and Large White pigs by RT-PCR

PCR primers were chosen to amplify cDNAs from exon 4 to the 3UTR. The RPL32 was chosen as an internal control for transcription levels. Tissue types are indicated on top of the figure. Sizes of the amplicons are indicated for bands that are specific to SLA-7 RNAs.

1.2.2.c- SLA-8

Primers suitable for amplification of cDNAs from exon 4 to the 3UTR were chosen for RT-PCRs (Table VI, primers SLA-8-e4_F/SLA-8-3UTR_R). As expected, a single band of 316 bp was found (Figure XXII), confirming the existence of a unique SLA-8 transcript, which corresponds to the full coding sequence of SLA-8 (Figure XVIII).

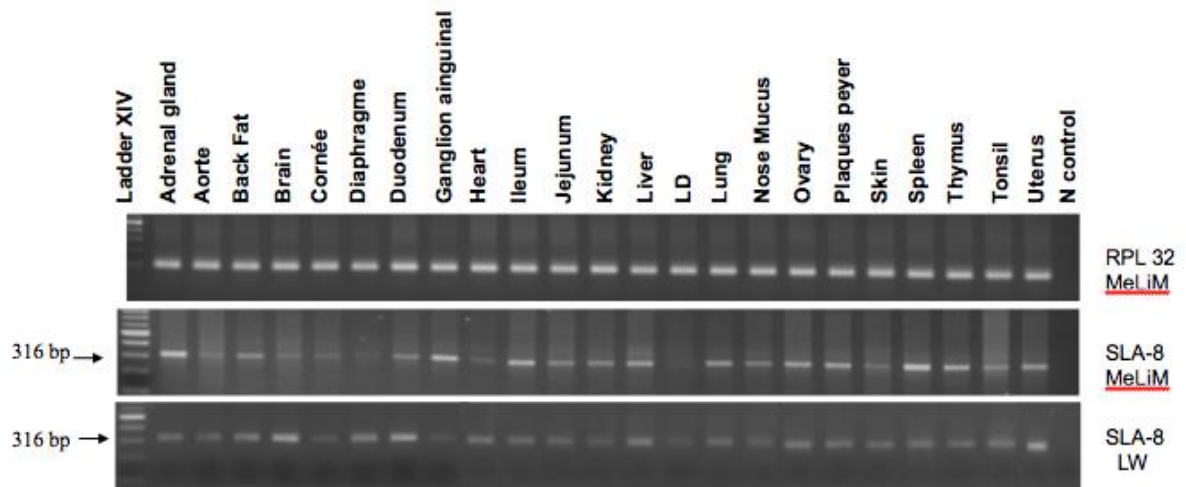


Figure XXII: Detection of SLA-8 RNAs in various tissues from MeLiM and Large White pigs by RT-PCR

PCR primers were chosen to amplify cDNAs from exon 4 to the 3UTR. The RPL32 was chosen as an internal control for transcription levels. Tissue types are indicated on top of the figure. A single band of 316 bp was found in all tissues from the two breeds.

2- SLA Ib protein

Addressing the role of SLA-Ib molecules is a major goal and we still need to produce tools to address this question. Actually, we could summarize the mains questions related to SLA-Ib proteins as follows: i)- do the SLA-Ib molecules have the ability to bind B2M and present peptide to cell surface as shown for HLA-G and HLA-E? ii)- are some SLA-Ib molecules mostly cytoplasmic as shown for HLA-F? As indicated previously, we still lack antibodies to efficiently work on these fundamental questions. In order to provide a start for such functional analyses, we have predicted the protein folding for each SLA-Ib splice variant. Moreover, we have set up transfection experiments and collaborated with a private Company to produce monoclonal antibodies that are still under testing.

2.1- Protein structure

Using bioinformatics analysis tools, SLA-6, -7 and -8 transcripts were translated to proteins and corresponding protein structures were predicted in order to analyse the peptide presentation ability of each SLA Ib transcript.

2.1.1- Predicted SLA-6 protein isoforms

The five SLA-6 variants (Figure XVI) were translated into proteins and the predicted 3D-structures are presented in Fig. XXIII.

The SLA-6 transcript that corresponds to the full coding sequence comprises 7 exons and encodes a protein with the canonical organisation of MHC class I proteins (Figure XXIII A and B). The full length SLA-6 protein isoform should have the ability to non covalently bind to B2M and the presence of alpha 1 and alpha 2 domains is suitable for peptide binding and

presentation on cell surface. These two properties are predicted but remain to be demonstrated.

The three RNA variants SLA-6-1, SLA-6-2 and SLA-6-3 lack an exon 3, which means that the encoded molecules do not have an alpha 2 domain. Therefore, these three isoforms cannot form a binding groove and should have no ability to present peptides. The three isoforms differ by the size of the cytoplasmic tail and the presence/absence of a transmembrane domain. Indeed, the isoform SLA-6-3 could be soluble whereas the other isoforms could be bound at the membrane.

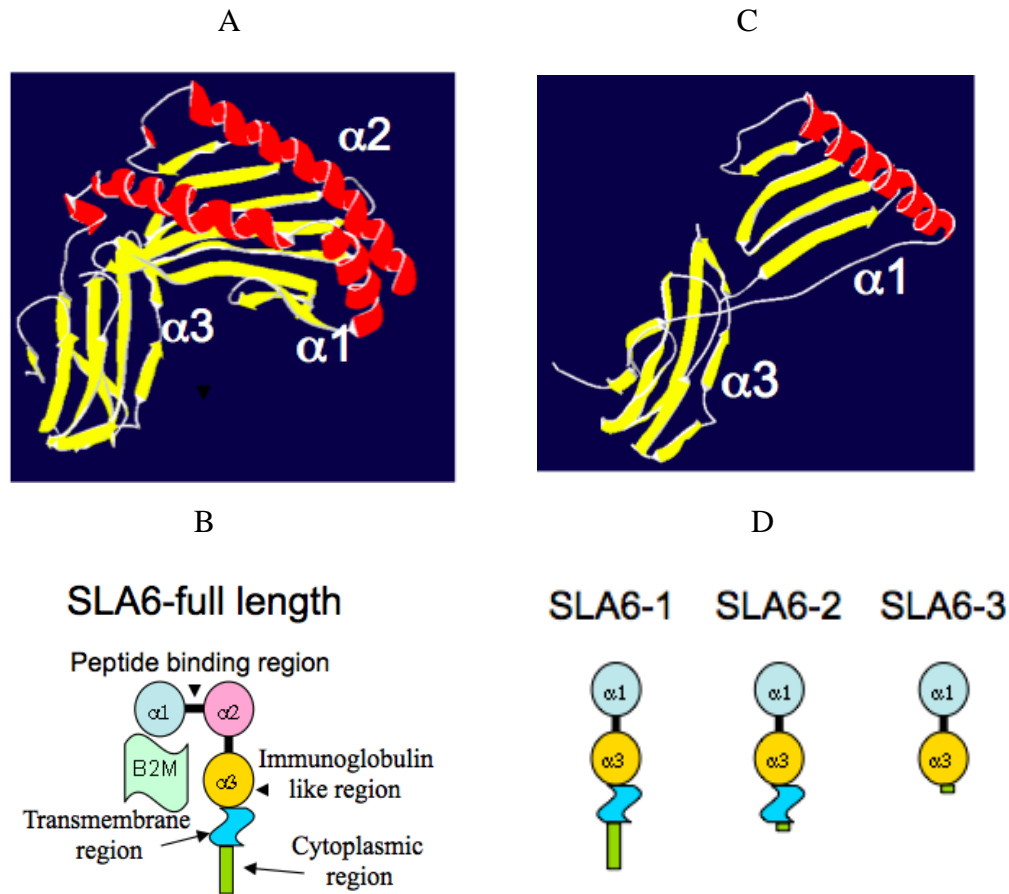


Figure XXIII: Prediction of SLA-6 protein encoded by the various splice variants

Putative folding of the protein part specific to alpha 1, 2 and 3 domains has been predicted for the full length SLA-6 isoform (A) and isoforms SLA-6-1, -2 and -3 missing the alpha 2 domain (C). Predictions were done with the CPHmodels 3.0 software. The functional domains of the protein isoforms are shown in B and D.

2.1.2- Predicted SLA-7 protein isoforms

Among the six SLA-7 transcripts found in our study (Figure XVII), only SLA-7 and SLA-7-1 encode molecules that are predicted to form a binding groove suitable for peptide presentation on the cell surface (Figure XXIV). The SLA-7 isoform has a molecular structure that is canonical for MHC class I molecules and suggests a good ability for peptide

presentation on cell membrane and a non covalent binding to B2M (Figures XXIV A and B). In contrast, the SLA-7-1 isoform harbours a shortened alpha 3 domain that may modify the ability to bind to B2M molecules (Figures XXIV C and D).

The transcripts SLA-7-2 and SLA-7-3 are spliced in exon 1 (22 missing nucleotides) and connect a short exon 1 of 42 nt to exon 2 (Figure XVII). The splice creates a premature termination codon early in exon 2. The encoded peptide would be 18 AA long and partially correspond to the leader peptide encoded by exon 1 removed during protein maturation.

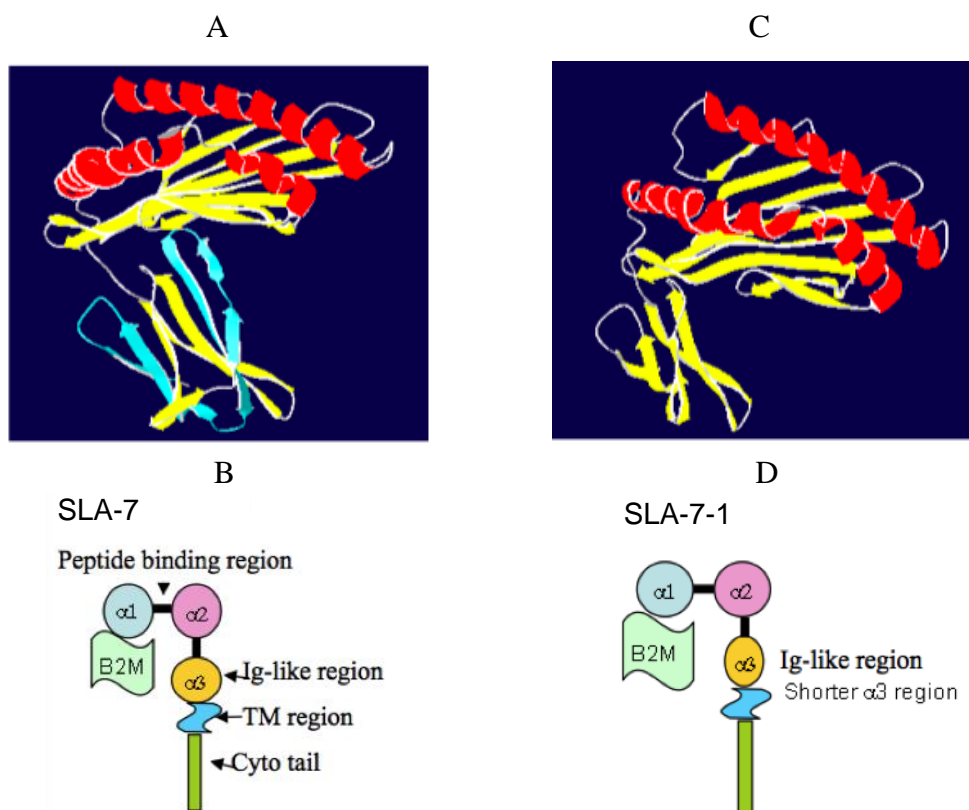


Figure XXIV: Prediction of SLA-7 protein isoforms encoded by the splice variants SLA-7 (A, B) and SLA-7-1 (C, D)

Putative folding of the protein part specific to alpha 1, 2 and 3 domains (A, C) has been predicted with the CPHmodels 3.0 software. The functional domains of the protein isoforms are shown in B and D.

2.1.3- Predicted SLA-8 protein

A single SLA-8 transcript was found (Figure XVIII) and the encoded protein has a structure that is canonical to MHC class I molecules (Figure XXV). The full length SLA-8 molecule is predicted to form a binding groove suitable for peptide presentation on cell surface and to non covalently bind to B2M.

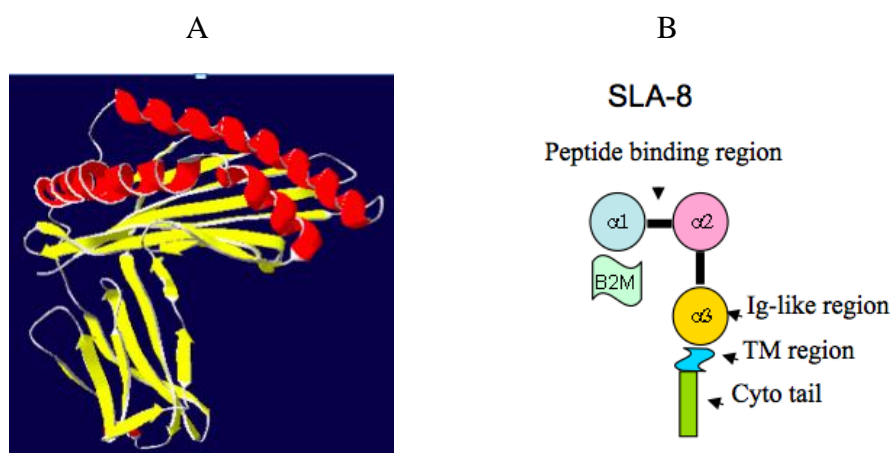


Figure XXV: Prediction of SLA-8 protein

Putative folding of the protein part specific to alpha 1, 2 and 3 domains (A) has been predicted with the CPHmodels 3.0 software. The functional domains of the protein are shown in B.

2.1.4- Conclusion on protein prediction

Our results confirm that the three SLA-Ib genes express full coding RNAs that encode molecules predicted to have a canonical molecular structure suitable for B2M binding and peptide presentation on cell surface (Chardon et al. 2001). SLA-6 and SLA-7 express additional RNA variants that encode putative isoforms with prediction for other properties, such as solubility (SLA-6-3), a possible modified binding to B2M (SLA-7-1), alternative activity of the cytoplasmic tail that is shortened in few SLA-6 isoforms (SLA-6-2, SLA-6-3).

In addition, the two full length SLA-7 molecules (SLA-7-001 corresponding to the RefSeq RNA and SLA-7 found in our study) differ in the length and sequence of cytoplasmic tail. By contrast to SLA-6 and SLA-7, a unique SLA-8 protein is predicted, suggesting less variation in its functions than for the various SLA-6 and -7 protein isoforms.

For further expression studies, we only focussed on expression of the full SLA-6, -7 and -8 molecules encoded by the RNAs referred to as SLA-6 (Figure XVI), SLA-7 (Figure XVII) and SLA-8 (Figure XVIII).

2.2- Protein expression

2.2.1- Experimental design of the experiments

Detection of SLA-Ib molecules is an important issue in order to go for functional analyses. In this study, we had two goals: i)- to design expression systems for transient and stable transfection of each SLA-Ib molecule, ii)- to produce monoclonal antibodies for specific detection of SLA-Ib molecules.

For antibody production, we collaborated with the P.A.R.I.S. Company.

For transfection experiments, a first aim was to detect the molecules on cell surface by FACS. We transiently transfected PK-15 cells with eukaryotic expression vectors suitable for expression of SLA-6, SLA-7 or SLA-8 molecules under the control of the cytomegalovirus (CMV) promoter. Transient transfections of PK-15 cells were carried out for the screening of monoclonal antibodies produced by the P.A.R.I.S. Company. The characterization of antibody specificities is still in progress.

A second much more challenging aim was to identify which of the three SLA-Ib molecules would be a good candidate as a homologous counterpart of HLA-E that is reported to present class Ia-derived peptides on cell surface. We set up transient and stable transfections in drosophila Schneider cells (SC2) known to have no MHC expression background and widely used to study expression of MHC molecules on cell surface. Since the cells grow at room temperature, the MHC molecules accumulate on surface due to a reduced traffic between cell surface and cytoplasm. For this purpose, we constructed expression vectors suitable for expression of B2M, SLA-1, SLA-6, SLA-7 or SLA-8 under the control of the drosophila actin 5 promoter (PAc5) (see Figure XII in Materials and Methods section).

We developed tests for detection of the expressed proteins by FACS and western blots. Most experiments with SC2 cells were performed before availability of monoclonal antibodies and we tried to use B2M antibodies suitable for detection of the molecules on cell surface where bound with MHC class I molecules. We will present the SC2-related experiments but could not demonstrate expression of the SLA-Ib molecules on cell surface with this design.

2.2.2- Construction of the expression vectors

2.2.2.a- Expression vectors for expression in mammalian cells

Expression vectors suitable for expression of SLA-6, SLA-7 or SLA-8 in mammalian cells were constructed by the In Cell Art Company (Nantes, France) in the frame of the antibody production by P.A.R.I.S. The full coding sequence inserts correspond to sequences from MeLiM pigs (see supplementary files SI and SII) and their structural specificities are summarized in table XIII.

2.2.2.b- Expression vectors for expression in *Drosophila* cells

Since SC2 cells do not express MHC and B2M molecules, our goal was to set up a system in which various expression combinations could be compared. If B2M is required for expression of SLA class Ib molecules on cell surface, co-expression of SLA-Ib and B2M is necessary, which means that detection of SLA-Ib molecules on cell surface is expected only by co-transfecting expression vectors for SLA-Ib and B2M. We also included a vector suitable for expression of SLA-1 molecules in order to have the possibility to co-express SLA-Ib and Ia molecules and test the hypothesis that SLA-Ib molecules may present leader peptides derived from class Ia molecules.

We constructed expression vectors suitable for expression of SLA-1, -6, -7, -8 (MeLiM pigs) and B2M proteins (see Figures XII and XIII). Insert sequences are presented in supplementary files SI and SII.

Table XIII: Characteristics of the inserts sub-cloned into expression vectors

Gene	Length (bp)	The first nt position ¹	Stop codon	Poly (A)	# exon
SLA-6	1221	-7	yes	no	7
SLA-7	1465	1	yes	no	8
SLA-8	1123	1	yes	no	8
SLA-1	1387	-35	yes	no	8
B2M	430	-25	yes	no	4

¹ This position refers to the initiation codon (ATG) for which the nucleotide A is numbered as the first one.

The inserts were initially cloned into a TA cloning kit before transfer into the pAc5.1/V5-His vector. Insert sizes and characteristics are presented in Table XIII and Figure XXVI.

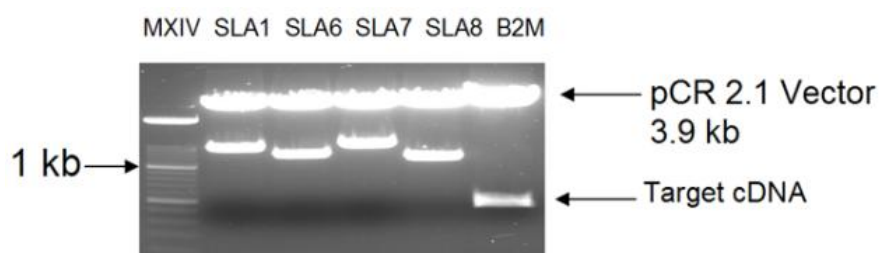


Figure XXVI: The expression vectors of SLA-6, -7, -8, -1, and B2M

Electrophoresis gel showing the inserts that were sub-cloned into the TA-cloning kit and further transferred into expression vectors by digesting with the restriction enzyme EcoR1.

2.2.3- Protein expression in drosophila cells

We carried out various experiments and report results on stable transfection of pAc5.1-B2M vector or stable co-transfection of pAc5.1-B2M with pAc5.1-SLA-1, SLA-6, SLA-7 or SLA-8 vectors. For stable transfection, the expression vectors containing the target genes to be expressed were co-transfected with the pCoBlast vector. Cell selection was carried out with a blasticidin -supplemented culture medium and resistant clones were pooled.

2.2.3.a- Detection of target RNAs in transfected cells

Total RNAs from transfected cells were extracted and RT-PCRs were carried out with gene specific primers (Table VII). B2M transcripts were detected by RT-PCR in lines stably transfected with pAc5.1-B2M or stably co-transfected with pAc5.1-B2M and either pAc5.1-SLA-1, pAc5.1-SLA-6, pAc5.1-SLA-7 or pAc5.1-SLA-8. We used the primers B2M-F and B2M-R (Table V) and the expected PCR product size was 420bp (Figure XXVII).

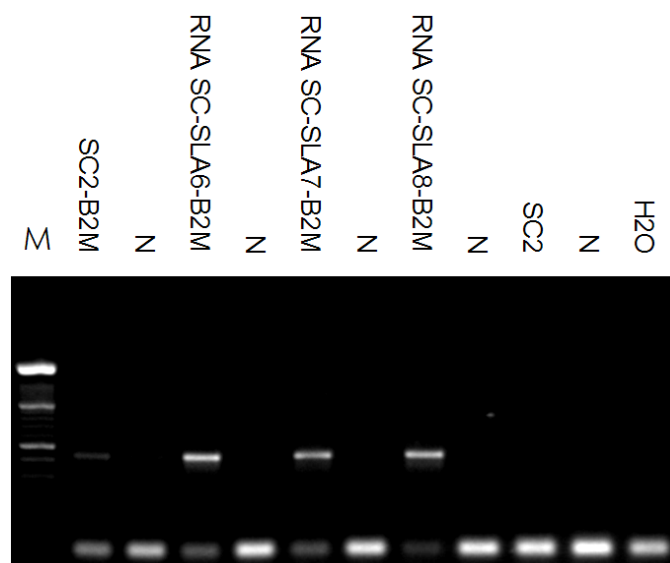


Figure XXVII: Detection of B2M transcripts by RT-PCR

B2M transcripts transfected with pAc5.1-B2M in S2 cells stably were detected by RT-PCR, including co-transfected with pAc5.1-B2M and pAc5.1-SLA-Ib vectors. N: negative RT controls. M: DNA molecular marker.

Our results clearly show that the B2M gene is efficiently transcribed in the four stably transfected cell lines that we have constructed as well as in the cell line co-transfected with pAc5.1-B2M and pAc5.1-SLA-1 vectors (data not shown).

2.2.3.b- Detection of target proteins in transfected cells

For protein detection, we do not have SLA-Ib specific antibodies. We used antibodies specific to B2M proteins that recognize the porcine B2M molecule. We could not detect the B2M protein by Western blot in stably transfected cells and the experiment was not repeated in order to more focus more on detection by flow cytometry (FACS).

For detection of SLA molecules on cell surface, we tried to use anti-B2M antibodies (Table VIII) and the porcine PK15 cells were chosen as positive control cells that constitutively

express the B2M molecules. FACS analysis (Figure XXVIII) clearly confirms that B2M is expressed on the surface of PK15 cells suggesting that the antibody recognizes the B2M protein that is bound to SLA class I molecules at the membrane. The non transfected S2 cells are negative as expected. No signal could be detected with cells transfected with pAc5.1-B2M vector, confirming that B2M cannot migrate to the cell surface if not bound to MHC class I molecules. Slight positive signals could be detected with cell lines co-transfected with pAc5.1-B2M and either pAc5.1-SLA-6, pAc5.1-SLA-7 or SLA-8, suggesting that a small cell sub-population from each cell line could express the SLA-Ib molecules at the membrane. These results are interesting but not really convincing and they need to be confirmed. These experiments will be reproduced as soon as anti-SLA-Ib monoclonal antibodies are available.

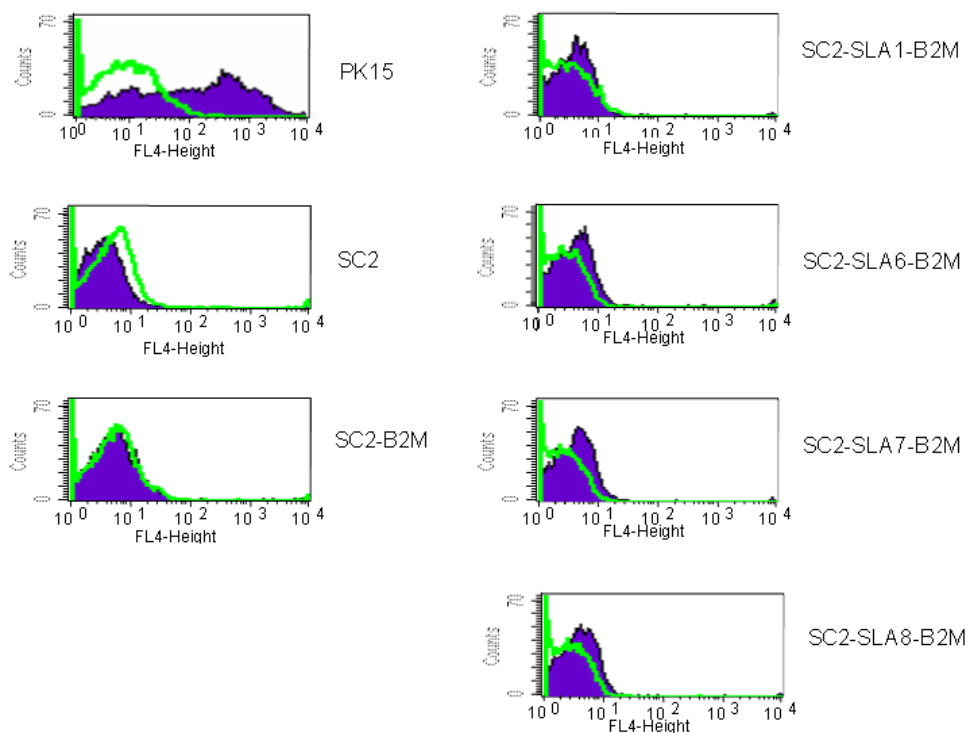


Figure XXVIII: Cell surface expression of B2M by FACS

Histograms obtained with fluorescence detected for isotype controls (green line) and fluorescence detected using anti-B2M antibodies (purple surface) was overlaid. Each cell line is indicated on the right hand of the histogram plots and refers to the cell names (PK-15 or SC2) and the genes that had been stably transfected (B2M, SLA-1, -6, -7 or -8).

2.2.4- Protein expression in PK15 cells and antibody screening

We set up transient transfection experiments in the porcine renal epithelial cell line PK15 to screen anti-SLA-Ib monoclonal antibodies produced by the P.A.R.I.S. Company. Mice were immunized either by a DNA vaccination strategy or a peptide strategy (see supplementary files SII and SIII) and data on SLA-Ib specificity of polyclonal sera from immunized mice were promising (see supplementary files SIII). We further tested monoclonal antibodies expected to be specific for SLA-6 and SLA-8.

PK15 cells were transfected with pVAX1-SLA6 or pVAX1-SLA8 vectors for 48 hours and the staining protocols were suitable to detect either the proteins in the cytoplasm by cell permeabilization or on cell surface (Figure XXIX). We tested 24 monoclonal sera for SLA-6 (Table XIV) and 23 for SLA-8 (Table XV).

Among the 24 sera expected to be specific for SLA-6 molecules, the three sera SLA6-2, -1, -3 mostly recognized molecules on surface of PVAX1-SLA6-transfected PK15 cells (no cell permeabilization) and the two sera SLA6-3 and -4 recognized molecules within the cytoplasm also (cell permeabilization). Other monoclonal antibodies recognized molecules in the cytoplasm as well as on cell surface but with a weaker intensity. Illustrations of the detection of SLA molecules on cell surface and in the cytoplasm of transfected PK15 cells using the sera SLA6-10 (negative), SLA6-1 (strongly positive) and SLA6-20 (slightly positive) are shown by dot plots and histogram plots (Figure XIX).

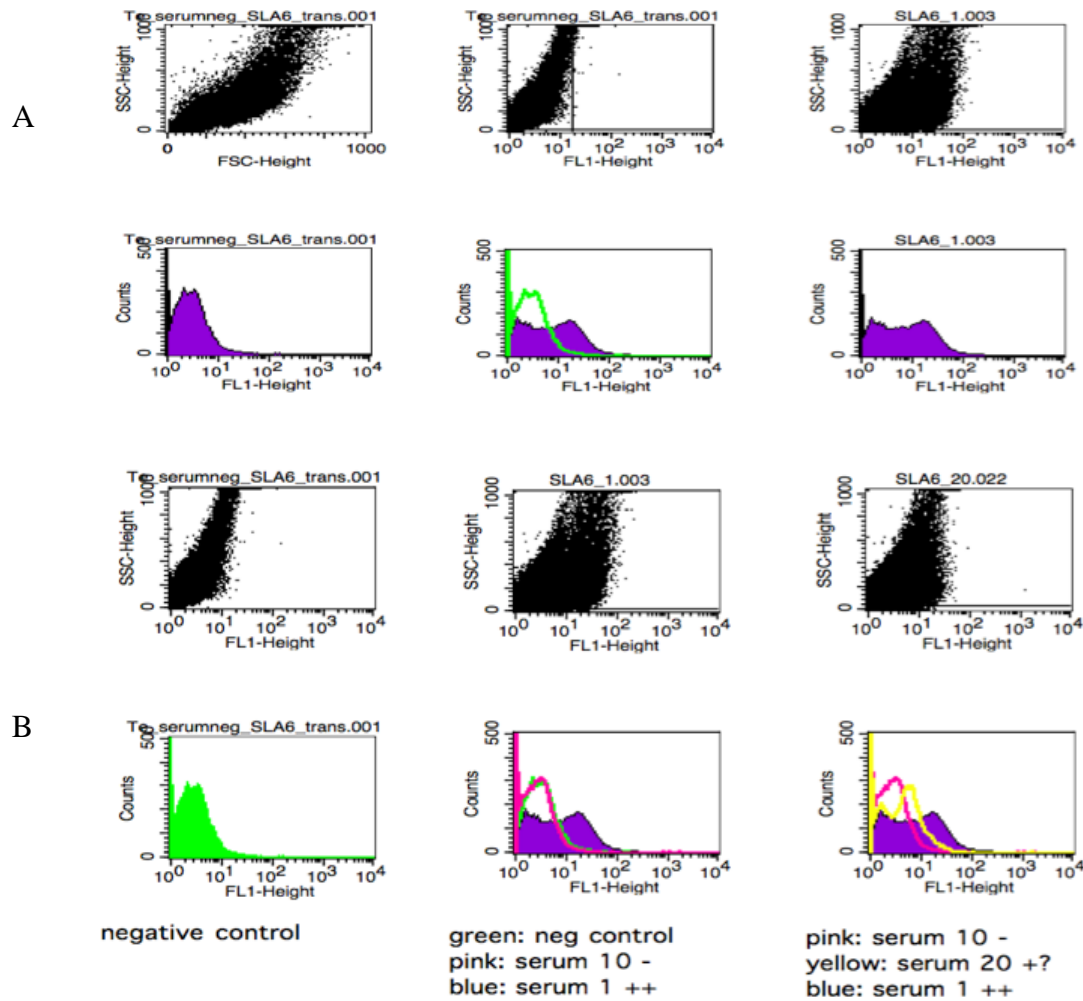


Figure XXIX: Detection of SLA molecules by FACS in the cytoplasm (A) or on surface (B) of PK15 cells transiently transfected with the vector pVAX1-SLA6

Cells were stained either with the SLA6-1 serum (histogram in blue) or the SLA6-10 serum (histogram in pink) or the SLA6-20 serum (histogram in yellow).

Among the 23 sera expected to be specific for SLA-8 molecules, the three sera SLA8-11, -12, -10 and -9 recognized SLA molecules in pVAX1-SLA8-transfected PK15 cells on cell surface. Illustrations of the detection of SLA molecules with the serum SLA8_2 (negative) and SLA8_11 (positive) are shown by dot plots and histogram plots (Figure XXX).

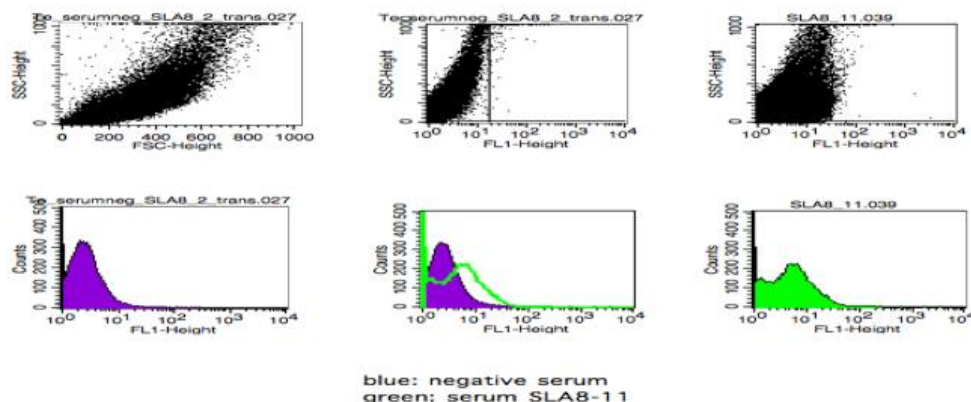


Figure XXX: Detection of SLA molecules by FACS on surface of PK15 cells transiently transfected with the vector pVAX1-SLA8

Cells were stained either with the SLA8_1 serum (histogram in blue) or the SLA8_11 serum (histogram in green).

We are currently testing cross-reactions of the antibodies against the various SLA class I molecules and at the moment, we cannot exclude that the produced monoclonal antibodies recognize SLA-Ia molecules instead of SLA-Ib molecules or recognize both types of molecules. Ongoing experiments show that mock-transfected cells express molecules that are recognized by the monoclonal antibodies selected as positive in our first screening. More experiments are required to assess the true specificity of the monoclonal antibodies under screening. We anticipate that we will have to test the protein specificity of the antibodies by transfecting the SC2 cells or mammalian cell lines known not to express SLA molecules such as the human HeLa cells.

Table XIV: Screening of anti-SLA-6 monoclonal antibodies by FACS in PK15 cells transiently transfected with the pVAX1-SLA6 vector

Sample ID	% Gated_UR ¹		Comment
	On cell surface	Intra cellular	
SLA6_2	24,38	26,42	Selection for further analyses
SLA6_1	21,71	21,75	
SLA6_3	10,22	23,61	
SLA6_22	4,87	3,67	
SLA6_19	4,74	2,22	
SLA6_4	4,43	14,57	
SLA6_16	2,68	1,57	
SLA6_20	2,62	1,12	
SLA6_6	2,57	1,99	
SLA6_24	2,31	2,81	
SLA6_8	2,15	3,13	
SLA6_18	2,13	1,8	
SLA6_17	2,03	1,37	
SLA6_5	1,81	1,34	Not selected
SLA6_13	1,75	2,27	
SLA6_21	1,53	1,46	
SLA6_14	1,32	2,92	
SLA6_7	1,19	1,15	
SLA6_15	0,92	1,66	
SLA6_11	0,74	3,1	
SLA6_23	0,71	1,67	
SLA6_9	0,68	1,69	
SLA6_12	0,48	1,72	
SLA6_10	0,35	0,83	
Mock_transfection_SLA6_2 ²	26,28	nd	Controls
Transfection_neg_serum ³	0,69	1.70	
Mock_transfection_SLA6_19 ⁴	nd	0.38	

¹ Percentage of positive events

² Mock-transfected PK15 cells stained with the SLA6_2 monoclonal antibody

³ Transfected PK15 cells stained with a control negative serum

⁴ Mock-transfected PK15 cells stained with the SLA6_19 serum

Table XV: Screening of anti-SLA-8 monoclonal antibodies by FACS in PK15 cells transiently transfected with the pVAX1-SLA8 vector

Sample ID	% Gated_UR ¹		Comment
	on cell surface	intra cellular	
SLA8_11	5,67	0,18	Selection for further analyses
SLA8_12	4,22	1.92	
SLA8_10	3,38	0,2	
SLA8_9	2,39	1.03	
SLA8_6	2,06	0,39	
SLA8_7	1,57	2.96	
SLA8_13	1,55	0,23	
SLA8_5	1,54	1.36	
SLA8_19	0,86	<0.1	Not selected
SLA8_18	0,73	<0.1	
SLA8_3	0,69	<0.1	
SLA8_2	0,68	0.70	
SLA8_17	0,66	<0.1	
SLA8_4	0,66	0,13	
SLA8_15	0,62	<0.1	
SLA8_1	0,51	<0.1	
SLA8_22	0,5	<0.1	
SLA8_23	0,49	<0.1	
SLA8_14	0,48	<0.1	
SLA8_16	0,46	<0.1	
SLA8_20	0,43	<0.1	
SLA8_8	0,39	<0.1	
SLA8_21	0,37	<0.1	
SLA8_21	0,35	<0.1	
mock_transfection_SLA8_2 ²	1,36	0.25	Negative controls
transfection_neg_serum ³	0,49	0.31	

¹ Percentage of positive events

² Mock-transfected PK15 cells stained with the SLA8_2 monoclonal antibody

³ Transfected PK15 cells stained with a control negative serum

3- SLA Ib polymorphism

As previously mentioned, very few SLA-Ib allelic sequences are available. One aim was to sequence SLA-Ib genes from various animals in order to verify that the polymorphism of SLA-Ib genes is limited. A second aim will be to characterize the SLA class I haplotypes by including allelic data from Ia and Ib genes in the future (not achieved now). For this work, the reference sequence corresponds to the haplotype Hp1a.0 and all mutations refer to this haplotype sequence.

3.1- cDNA sequencing: identification of coding SNPs (cSNPs)

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variations. The SNPs located on coding regions of genes are referenced as coding SNPs (cSNPs). They can be identified via cDNA sequencing.

3.1.1- SLA-6

In our SLA Ib transcription study, five SLA-6 transcripts were found in MeLiM pigs. By aligning all transcripts, four SNPs including one cSNP were identified (Table XVI). The cSNP occurred at the 813 position within exon 4. This missense mutation from G to A (G813A) induced an amino acid change from Arginine to Glycine. The three other SNPs that were found in the 3'UTR region were not cSNP, but might function as regulatory SNPs (rSNPs) to adjust gene expression. The SNP located on 1116 is a transversion from C to T (C1116T) and created a CdxA binding site. CTATAATT is a binding site for the mammalian caudal-type homeobox domain transcription factor CdxA, which was found in chicken and

could enhance the transcriptional level of a target gene (Margalit et al. 1993). The insert of T at 1202 also created a new CdxA binding site. (Table XVI).

Table XVI: SLA-6 Coding SNP Position

Transcript \ Position	Exon4	3'UTR
SLA-6 (1221bp)		
SLA-6-1 (945bp)	G813A	
SLA-6-2 (1812bp)		
SLA-6-3 (912bp)		A1181C
SLA-6-4 (811bp)		C1116T A1181C -1202T

3.1.2- SLA-7

Five full length transcripts of SLA-7 were cloned from MeLiM pigs. Alignment analysis of those five cDNA sequences identified 9 cSNPs in exon 1, exon 2, exon 3 and exon 4 (Table XVII). The SNPs A11G in exon1 and G350A in exon 3 were silent mutations. In the transcript SLA-7 (length = 1465 nt), the SNP A to G occurred at position 499 within exon3, and the encoded amino acid was changed from Arginine to Lysine. In the transcript SLA-7-1 (length=1366 nt), the mutation from A to G at position 179 induced an amino acid change from Asparagine to Serine. On the same transcript, the T to C mutation at position 437 induced a Valine to Alanine change. The mutation A824G in exon 4 induced a change from Valine to Alanine. Although three SNPs (A119G, A500G, and T578C) were found in exon2 and exon3 of the transcript SLA-7-3 (length=1101 nt), these mutations could not be related to any amino acid change because the deletion in exon1 creates a premature stop codon. However, these mutations could create amino acid modifications if present in gene alleles coding full length transcripts.

Table XVII: SLA-7 Coding SNP position

Position Transcript	Exon1	Exon2	Exon3	Exon4
SLA7-1465			G350A, A499G	
SLA7-1366		A179G	T437C	A824G
SLA7-1443a	A11G			
SLA7-1101	A11G	A119G	A500G, T578C	
SLA7-1443b	A11G			

3.1.3- SLA-8

For SLA-8, a unique transcript was found and the cDNA sequence was similar to the SLA-8 reference sequence. These first results show that SLA-7 seems more polymorphic than the two other SLA-Ib genes.

3.2- Characterization of the nucleotide polymorphism at the whole gene level

3.2.1- Experimental design

In order to characterize cSNPs and regulatory SNPs (rSNPs), we set up long PCRs for further sub-cloning and sequencing. PCR primers were designed to amplify each SLA Ib gene, from upstream of the promoter to downstream of the polyadenylation signal. Internal primers were chosen on both genomic DNA strands to apply a primer-walking sequencing method. The positions of primers are shown in Figure XXXI. Totally, including PCR and internal sequencing, 14 pairs of primers were designed for SLA-6 and SLA-7 separately, and 13 pairs of primers for SLA-8.

SLA Ib genomic amplification primers

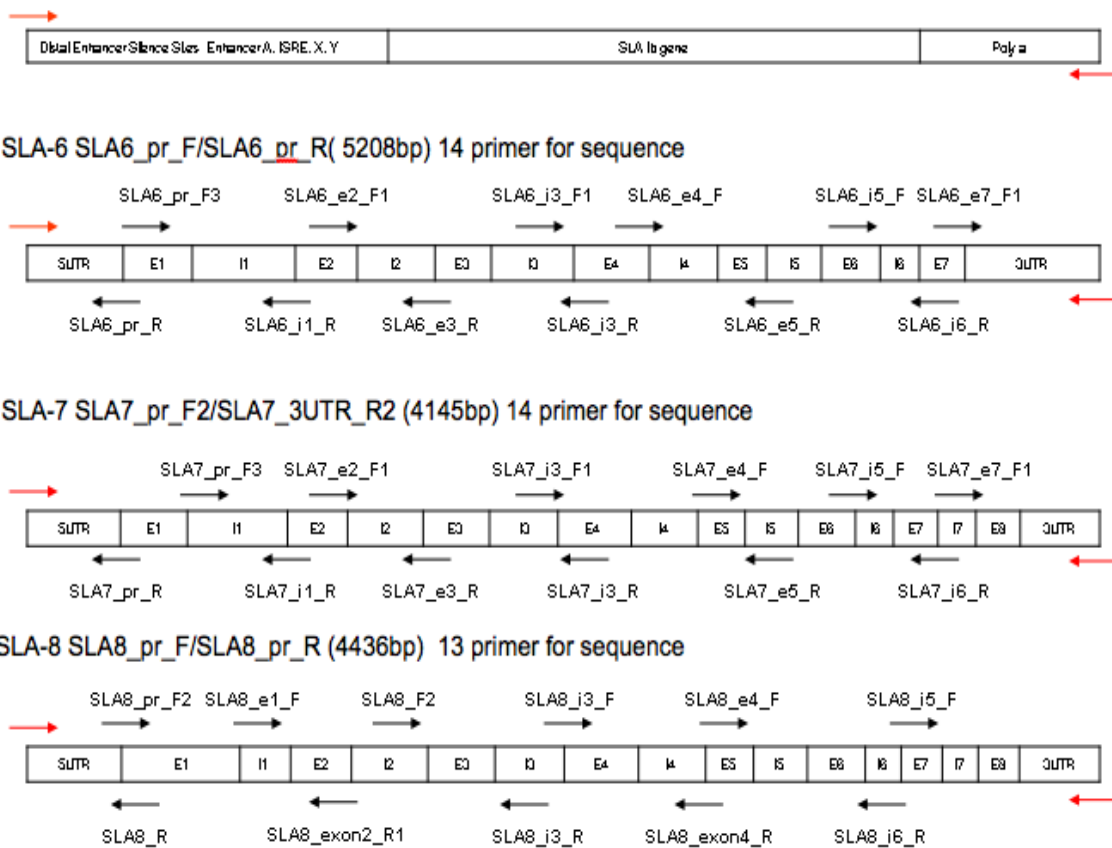


Figure XXXI: The primer positions used in primer walking sequencing

The red arrows represent the primer used to amplify the SLA Ib genes, and the black arrows represent the internal sequencing primers chosen in both orientations.

The amplicons were 5208 bp, 4145 bp, and 4436 bp for SLA-6, -7 and -8, respectively (Figure XXXII). These PCR products were sub-cloned into the vector XL TOPO and the clones harbouring the expected insert size were selected for sequencing.

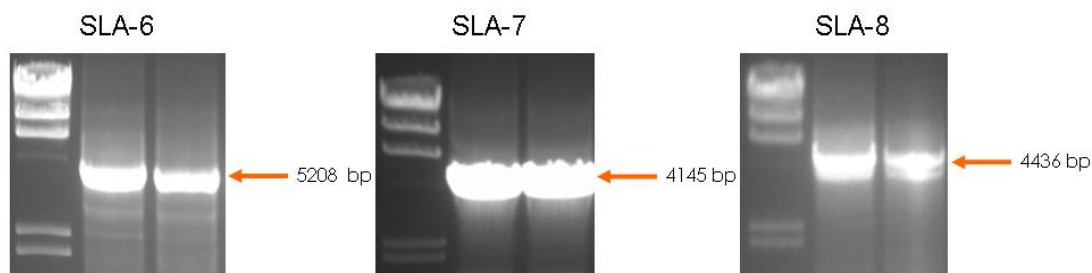


Figure XXXII: Gel electrophoresis showing amplification of SLA-6, -7 and -8 genes by long range PCRs

The amplicon size is indicated for each gene.

In this study, the three SLA Ib genes were amplified from six MeLiM pigs (5 MeLiM founders and 1 MeLiM from the herd) and one Large White pig harbouring the Hp1a.0 SLA haplotype.

As shown in figure XXXII, we could amplify three SLA-Ib genes. Unfortunately, we could not easily subclone SLA-6 amplicons and polymorphism analyses were only pursued for SLA-7 and SLA-8 genes. The total number of sequenced clones is indicated in Table XVIII with the corresponding animal. For SLA-7, five or six clones were sequenced per animal leading to a total of sequenced 39 clones. For SLA-8, one, four or five clones were sequenced per animal, resulting in a total of 10 sequenced clones.

Table XVIII: Number of clones selected per animal for sequencing SLA-7 and SLA-8 genes Clone information

Animal Gene	MeLiM B52	MeLiM C284	MeLiM C321	MeLiM F206	MeLiM F213	MeLiM 486	Large white 5599
SLA-7	6 clones	5 clones	5 clones	5 clones	6 clones	6 clones	6 clones
SLA-8			1 clones		4 clones	4 clones	5 clones

3.2.2- SLA-7 polymorphism

For SLA-7, 90 SNPs were identified (Table XIX) from six MeLiM pigs and one Large White pig (Table XVIII). We identified 14 SNPs in the 5'UTR and 16 SNPs in the 3'UTR. In addition, a deletion polymorphism from 3974 to 3979 (6 nucleotides) was found in the 3'UTR. This 6 nucleotides deletion was detected in 10 clones from 3 MeLiM pigs. 18 SNPs were identified in six among the eight SLA-7 exons: 3 SNPs in exon 2, 3 SNPs in exon 3, 7 SNPs in exon 4, 1 SNP in exon 5, 1 SNP in exon 6 and 3 SNPs in exon 7. No SNP was found in exons 1 and 8. 42 SNPs were identified in intronic sequences: 12 SNPs in intron1, 5 SNPs in intron 2, 17 SNPs in intron 3 and 8 SNPs in intron 5. No SNP could be detected in introns 4, 6 and 7 (Table XIX).

Table XIX: SNP positions on SLA-7 gene

Position	SNP	Number					Location
		A	C	G	T	-	
200	C > T		26		13		5'UTR
218	A > G	38		1			5'UTR
233	G > C		4	35			5'UTR
257	T > A	1			38		5'UTR
281	G > A	4		35			5'UTR
297	C > T		31		8		5'UTR
403	G > A	17		22			5'UTR
415	G > A	8		31			5'UTR
416	T > C		9		30		5'UTR
476	G > A	4		35			5'UTR
562	T > C		22		17		5'UTR
563	C > T		30		9		5'UTR
576	A > G	38		1			5'UTR
691	G > A	1		38			5'UTR
885	G > -			30		9	Intron1
886	T > -				30	9	Intron1
887	G > -			30		9	Intron1
888	G > -			30		9	Intron1
890	G > A	18		21			Intron1
929	T > C		1		38		Intron1
941	C > T		30		9		Intron1
953	A > C	21	18				Intron1
970	C > T		21		18		Intron1
1033	T > C		1		38		Intron1
1076	T > C		18		21		Intron1
1155	C > T		22		17		Intron1
1242	T > C		1		38		Exon2
1388	A > G	38		1			Exon2
1398	T > C		1		38		Exon2
1453	T > C		1		38		Intron2
1535	A > G	21		18			Intron2
1540	G > T			21	18		Intron2
1563	C > G		20	19			Intron2
1656	G > A	19		20			Intron2
1690	C > T		20		19		Exon3
1743	G > A	16		23			Exon3
1770	A > G	38		1			Exon3

Position	SNP	Number					Location
		A	C	G	T	-	
1918	A > G	38		1			Intron3
1926	A > G	38		1			Intron3
2048	A > -	30				9	Intron3
2049	A > -	30				9	Intron3
2050	G > -			30		9	Intron3
2051	G > -			30		9	Intron3
2052	A > -	30				9	Intron3
2053	T > -				30	9	Intron3
2054	C > -		30			9	Intron3
2055	T > -				30	9	Intron3
2056	A > -	30				9	Intron3
2057	G > -			30		9	Intron3
2058	G > -			29		10	Intron3
2161	T > C		1		38		Intron3
2294	A > G	35		4			Intron3
2378	A > G	38		1			Intron3
2448	A > C	23	16				Intron3
2572	C > A	18	21				Exon4
2579	T > C		1		38		Exon4
2620	T > C		1		38		Exon4
2719	C > G		23	16			Exon4
2735	T > C		9		30		Exon4
2738	C > T		38		1		Exon4
2818	G > A	4		35			Exon4
2934	G > A	18		21			Exon5
3065	T > C		1		38		Intron5
3235	C > -		30			9	Intron5
3236	A > -	30				9	Intron5
3243	C > T		38		1		Intron5
3317	G > A	9		30			Intron5
3343	C > G		30	9			Intron5
3405	T > C		1		38		Intron5
3406	T > A	1			38		Intron5
3484	T > C		1		38		Exon6
3601	C > T		24		15		Exon7
3718	C > T		35		4		Exon7
3742	G > -			30		9	Exon7
3746	G > A	1		38			3'UTR
3850	T > C		1		38		3'UTR
3881	C > A	19	20				3'UTR
3893	C > A	1	38				3'UTR

Position	SNP	Number					Location
		A	C	G	T	-	
3894	C > G		38	1			3'UTR
3927	C > T		36		3		3'UTR
3959	G > T			38	1		3'UTR
3974	T > -				29	10	3'UTR
3975	T > -				29	10	3'UTR
3976	A > -	29				10	3'UTR
3977	A > -	29				10	3'UTR
3978	G > -			29		10	3'UTR
3979	G > -			29		10	3'UTR
3980	A > G	35		4			3'UTR
4102	G > A	3		35			3'UTR
4120	G > A	9		29			3'UTR

3.2.3- SLA-8 polymorphism

For SLA-8, 60 SNPs were identified from three MeLiM pigs and one Large White pig (Tables XVIII and XX). But contrast to SLA-7, only four SNPs were found in coding sequences, suggesting that SLA-8 is less polymorphic than SLA-7. Three SNPs corresponded to a mutation from A to G in exon 2 (A1368G), exon 3 (A1953G), and exon 4 (A2823G). The fourth SNP was a mutation from C to T at position 2069 in exon 3. 21 and 2 SNPs were found in the 5'UTR and 3'UTR, respectively. In the 5'UTR, 13 of these SNPs were deletion polymorphisms. In addition, 33 SNPs were identified within intronic sequences. Intron 4 harbored the highest number of SNPs (11 SNPs) followed by intron 5 (10 SNPs), intron 7 (4 SNPs), intron 3 (3 SNPs), introns 1 and 3 (2 SNPs in each intron), and finally intron 6 (1 SNP). Within intron 4, the 11 SNPs corresponded to a two deletions covering 5 and 6 SNPs (Table XX).

Table: XX: SNP position on SLA-8 gene

Animal	486	486	486	486	C321	F213	F213	F213	F213	LW	LW	LW	LW	LW			
Clone No	2	3	6	7	3	1	2	3	4	1	2	3	4	5	Ref		
Position																Summary:	Position
135	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	14A,1G	5UTR
273	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	14-,1T	5UTR
274	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	13-,2T	5UTR
275	-	T	T	T	T	-	-	-	-	T	T	-	T	T	T	9T,6-	5UTR
276	T	T	T	T	T	-	-	-	-	T	T	-	T	T	T	10T,5-	5UTR
277	T	T	T	T	T	-	-	-	-	T	T	T	T	T	T	11T,4-	5UTR
295	T	T	T	T	T	T	-	T	-	T	T	T	T	T	T	13T,2-	5UTR
296	T	T	T	T	T	C	C	C	C	T	T	T	T	T	T	11T,4C	5UTR
301	T	T	T	T	-	T	T	T	T	-	-	T	T	T	T	12T,3-	5UTR
302	T	T	T	T	-	T	T	T	T	-	-	T	T	T	T	12T,3-	5UTR
303	-	-	T	T	-	-	-	-	-	-	-	-	-	T	T	11-,4T	5UTR
304	-	-	T	T	-	-	-	-	-	-	-	-	-	T	T	11-,4T	5UTR
364	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	14G,1A	5UTR
406	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	14A,1G	5UTR
524	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	14C,1T	5UTR
570	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	14C,1T	5UTR
579	-	-	-	A	A	-	-	-	-	A	A	-	A	-	A	9-,6A	5UTR
580	-	-	A	A	A	A	A	A	A	A	A	A	A	A	A	13A,2-	5UTR
591	A	A	A	A	-	-	A	A	-	-	-	A	-	A	-	8A,7-	5UTR
887	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	14T,1C	5UTR
956	T	T	T	T	T	T	A	T	A	T	T	T	T	T	T	13T,2A	5UTR
1127	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1G	intron1
1150	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	14T,1C	intron1
1368	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	14A,1G	exon2
1663	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	14G,1A	intron2
1700	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	14G,1A	intron2
1953	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1G	exon3
2069	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	14C,1T	exon3
2338	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	14A,1G	intron3
2601	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	14T,1C	intron3
2675	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1G	intron3
2823	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	14A,1G	exon4
3041	-	-	-	-	-	C	C	C	C	-	-	-	-	-	-	11-,4C	intron4
3042	-	-	-	-	-	C	C	C	C	-	-	-	-	-	-	11-,4C	intron4
3043	-	-	-	-	-	T	T	T	T	-	-	-	-	-	-	11-,4T	intron4
3044	-	-	-	-	-	T	T	T	T	-	-	-	-	-	-	11-,4T	intron4
3045	-	-	-	-	-	C	C	C	C	-	-	-	-	-	-	11-,4C	intron4
3046	-	-	-	-	-	G	G	G	G	-	-	-	-	-	-	11-,4G	intron4
3047	-	-	-	-	-	G	G	G	G	-	-	-	-	-	-	11-,4G	intron4
3048	-	-	-	-	-	C	C	C	C	-	-	-	-	-	-	11-,4C	intron4
3049	-	-	-	-	-	G	G	G	G	-	-	-	-	-	-	11-,4G	intron4
3050	-	-	-	-	-	G	G	G	G	-	-	-	-	-	-	11-,4G	intron4
3051	-	-	-	-	-	G	G	G	G	-	-	-	-	-	-	11-,4G	intron4

Animal	486	486	486	486	C321	F213	F213	F213	F213	LW	LW	LW	LW	LW			
Clone No	2	3	6	7	3	1	2	3	4	1	2	3	4	5	Ref		
Position																Summary:	Position
3203	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1-	intron5
3211	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1C	intron5
3236	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	14A,1G	intron5
3242	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	14T,1C	intron5
3275	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	14A,1G	intron5
3290	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	14T,1C	intron5
3474	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	14C,1T	intron5
3498	T	T	T	T	T	C	C	C	T	T	T	T	T	T	T	12T,3C	intron5
3542	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	14G,1A	intron5
3570	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	14T,1C	intron5
3694	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	14G,1A	intron6
3860	A	A	A	A	A	C	C	C	C	A	A	A	A	A	A	4C,11A	intron7
3912	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	14A,1G	intron7
3925	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	11A,4G	intron7
4005	C	C	C	C	C	T	T	T	C	C	C	C	C	C	C	12C,3T	intron7
4449	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	13G,1T	3UTR
Total # differences	10	4	9	5	5	19	22	19	19	7	6	7	4	7	6	149	3UTR

3.3- Copy Number Variation (CNV) of SLA Ib genes

CNVs have been reported in for MHC genes in various species, including the pig (Tanaka-Matsuda et al., 2009). By sequencing various clones specific for SLA-7 and SLA-8 in MeLiM pigs, we identified a much higher polymorphism of SLA-7 gene than SLA-8 gene. In addition, the number of SNPs within SLA-7 clones could not be explained by only two alleles of the same gene per animal. These results strongly suggested that MeLiM animals could harbour more than one copy of the SLA-7 gene in their genomes. The drawing of SLA-7 SNP haplotypes together with pyrosequencing-based experiments confirmed this hypothesis.

3.3.1- SLA-7 SNP haplotypes

Haplotypes were drawn for the MeLiM pig 484 using genomic polymorphism data detected by sequencing 21 amplicons that spanned a segment from intron 3 to 3' UTR. 15 SNPs were identified and five haplotypes could be characterized (Fig. XXXIII), confirming that more than one SLA-7 copy exist in the genome of the MeLiM pig 484.

# Hap	# Clone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SLA7-Hap-01	484-R2-12	A	G	T	C	G	T	G	C	T	T	G	C	A	A	G
	484-R2-19	A	G	T	C	G	T	G	C	T	T	G	C	A	A	G
	484-R2-15	A	G	T	C	G	T	G	C	T	T	G	C	A	A	G
	484-R2-16	A	G	T	C	G	T	G	C	T	T	G	C	A	A	G
SLA7-Hap-02	484-R2-33	A	G	T	C	G	T	G	C	T	T	G	C	A	A	G
SLA7-Hap-03	484-R2-14	A	C	C	C	A	T	A	G	T	C	-	A	A	A	A
	484-R2-29	A	C	C	C	A	T	A	G	T	C	-	A	A	A	A
	484-R2-32	A	C	C	C	A	T	A	G	T	C	-	A	A	A	A
	484-R2-02	A	C	C	C	A	T	A	G	T	C	-	A	A	A	A
SLA7-Hap-04	484-R2-26	A	C	C	C	A	T	A	G	T	C	-	A	A	G	A
	484-R2-01	A	C	C	C	A	T	A	G	T	C	-	A	A	G	A
SLA7-Hap-05	484-R2-28	G	C	C	A	A	C	A	C	T	T	G	C	A	A	G

484-R2-01 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-26 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-12 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-19 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-28 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-16 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...
 484-R2-33 TGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGG...

Figure XXXIII: Different haplotypes of SLA-7 identified in the MeLiM pig 484

3.3.2- SNP quantification by pyrosequencing

Pyrosequencing has been shown to be a straightforward method to quantify SNP alleles at polymorphic positions. Indeed, an imbalanced quantification of SNPs agrees with the presence of more than one copy of the amplified sequence in the genome. We applied this technique to look for CNVs in MeLiM pigs and further make sure that more than one SLA-7 copy exists in MeLiM pigs.

For the pyrosequencing reaction and SNP detection a specific position, three primers are required: two PCR primers (forward primer and reverse primer), and one sequencing primer. Three pairs of specific PCR primers were designed according to SLA-7 genomic sequencing results (Fig XXXIV).

A first pair of primer sequences was selected from position 1509 to 1564 within intron 2 (Fig. XXXIV-A). With this primer combination, the 2 SNPs A1535G and G1540T could be genotyped. SLA7_1535_F and SLA7_1535_RB primers were used to amplify the target sequence. The primer SLA7_1535_RB was labeled with biotin. A second pair of primers was designed to test the SNP at position 3601 (C3601T) in exon7 (Fig. XXXIV-B). SLA7_3601_F and SLA7_3601_RB were used to amplify the target region. A third pair of primers matched insert/deletion polymorphism located from 3974 to 3979 in the 3'UTR (Fig. XXXIV-C). The primers SLA7_3974_F and SLA7_3974_RB stood for amplification primers and SLA7_3974_S for the sequencing primer.



Position 3601 (SNP)

```

        3570      3580      3590      3600      3610      3620      3630      3640
    .....|.....|.....|.....|.....|.....|.....|.....|
    ---SLA7_3601 F--->
    GGAGCAGCCGTAGCGAGAA
              ----SLA7_3601 S---->
                CGAGAACTCTGGTGTATGT
SLA7-B52-1   GGAGCAGCCGTAGCGAGAACTCTGGTGTATGTTTGATGCCTTTCAAAGCTGAGACCCCTGGAGGCCCTGAGTGGGACAGC
SLA7-B52-2   GGAGCAGCCGTAGCGAGAACTCTGGTGTATGTTTCGATGCCTTTCAAAGCTGAGACCCCTGGAGGCCCTGAGTGGGACAGC
SLA7-B52-3   GGAGCAGCCGTAGCGAGAACTCTGGTGTATTGTCGATGCCTTTCAAAGCTGAGACCCCTGGAGGCCCTGAGTGGGACAGC

        3650      3660      3670      3680
    .....|.....|.....|.....|.....|.....|.....|.....|
SLA7-B52-1   GGTTGGGGTAGAGGGGACACAACAGGGTTATGGGGGT
SLA7-B52-2   GGTTGGGGTAGAGGGGACACAACAGGGTTATGGGGGT
SLA7-B52-3   GGTTGGGGTAGAGGGGACACAACAGGGTTATGGGGGT
                  CTGTGTTGCCCAATACCCCCA-biot
                   <-----SLA7_3601 RB-----

```

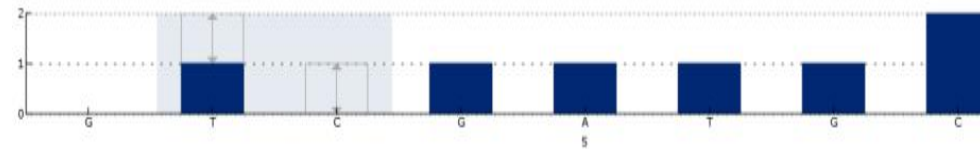
Primers for PCR :

SLA7_3601_F GGAGCAGCCGTAGCGAGAA
SLA7_3601_RB ACCCCATAACCCTGTTGTGTC

PCR length : 117 pb

Primer for sequencing :
SLA7_3601_S CGAGAACTCTGGTGTATGT

Dispensation order : GTCGATGC



C

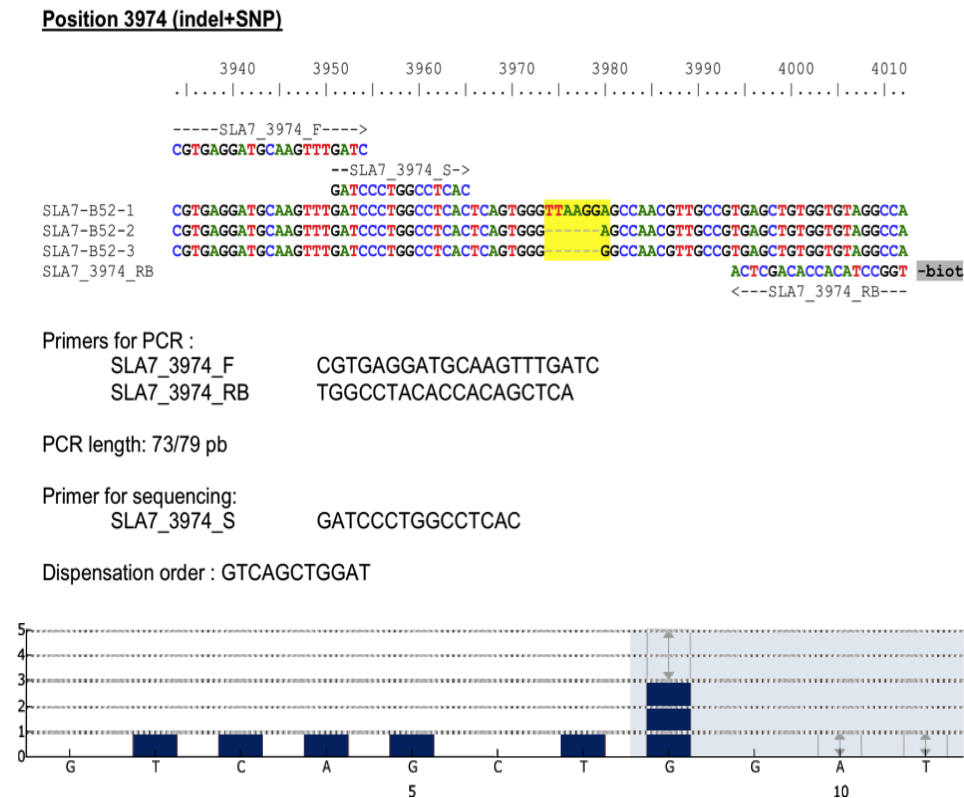


Figure XXXIV: Primer design for pyrosequencing

Primers were designed to genotypes the SNPs A1535G and G1540T (A), C3601T (B), and an indel at position 3274-3279. The sequence multi-alignments showing the polymorphism to detect are presented on top of each figure. The PCR and sequencing primers are indicated together with the amplicon length and the nucleotide dispensation order for pyrosequencing.

Using the three groups of primer pairs, SNP1535, SNP1540, SNP3601 and deletion 3974 were detected in four MeLiM founder pigs (IDs B52, C284, C321 and F206), 2 MeLiM offspring pigs (IDs 485 and 486), and a Large White pig with the Hp1a.0 haplotype. Pyrosequencing results showed that variability was found at positions SNP1535, SNP1540 and SNP3601. However no variability was identified for the deletion at position 3974.

Figure XXXV showed an example of pyrosequencing results with imbalanced quantification of SNPs at position SNP1535 and SNP1540 from a MeLiM pig (ID 485). For these two SNPs, the sequence with RGATCKCAGCTGATAGATCTCTTGGG was targeted sequence. Here R was A and G, and K was G and T. The Pyrosequencing trace of SNP1535 showed it included A/A and A/G two different alleles with the ratio 3:2 at this locus, and SNP1540 included G/G and G/T alleles with ratio 1:1. The summary result of allele frequency of SNPs was given in the Table XXI. And the further relative frequency of the paralogous sequences and ratios of different alleles were listed in Table XXII and XXIII.

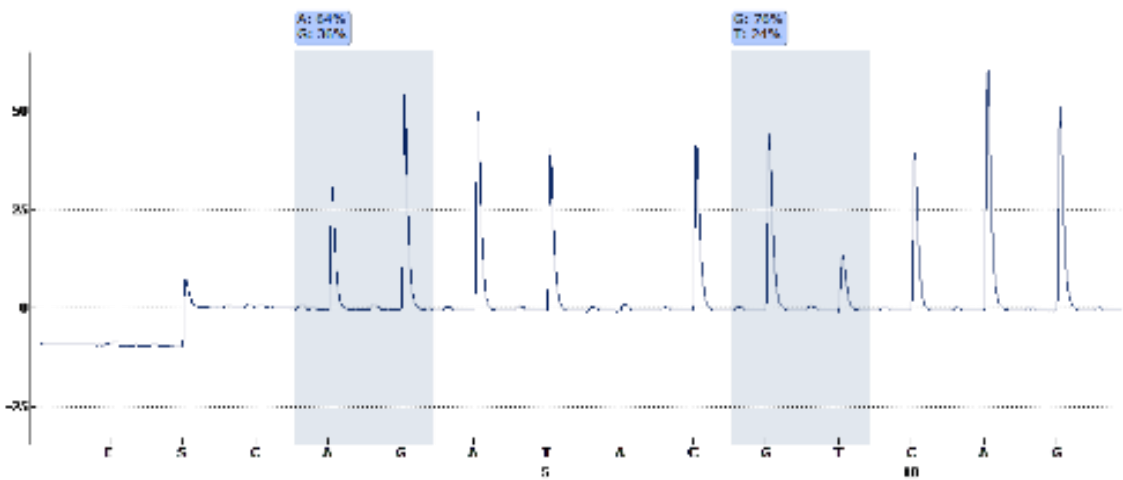


Figure XXXV: Pyrosequencing result for SNP1535 and SNP1540 from MeLiM pig

Tableau XXI: Allele frequency of SNPs at positions 3601, 1535 and 1540 by pyrosequencing

Assay	Sample ID	Var. Pos.	A (%)	C (%)	G (%)	T (%)
SNP3601	B52-3601	Y		50.19		49.81
	C284-3601	Y		50.1		49.9
	C321-3601	Y		100		0
	F206-3601	Y		99.37		0.63
	485-3601	Y		54.75		45.25
	486-3601	Y		51.14		48.86
	LW5599-3601	Y		1.56		98.44
SNP1535	B52-1535	R	64.32		35.68	
	C284-1535	R	56.41		43.59	
	C321-1535	R	61.79		38.21	
	F206-1535	R	84.61		15.39	
	485-1535	R	69.15		30.85	
	486-1535	R	59.65		40.35	
	LW5599-1535	R	100		0	
SNP1540	B52-1540	K			76.41	23.59
	C284-1540	K			70.38	29.62
	C321-1540	K			75.7	24.3
	F206-1540	K			88.25	11.75
	485-1540	K			77.3	22.7
	486-1540	K			72.91	27.09
	LW5599-1540	K			97.33	2.67

Among seven pigs, only the LW5599 Large White pig was T/T genotype. Two C/C animals were the MeLiM founder pigs (C321 and F206). And another four MeLiM pigs (B52, C284, 485, and 486) were determined with C/T genotype.

In contrast to SNP genotypes at position 3601, imbalanced genotypes were detected at positions SNP1535 and SNP1540 (table XXI) and further quantification assays using the Pyrosequencing AQ software were in agreement with various copy numbers (tables XXII and XXIII). As summarized in Table XXII, the Large White pig LW5599 harbored an A/A genotype and the MeLiM pig C284 an A/G genotype at position 1535. In contrast, both A/A and A/G genotypes were found in other MeLiM animals, meaning that more than one copy of SLA-7 exist in these animals. Similarly, the genotype of the Large White pig (LW5599) was G/G at position 1540 but two genotypes G/G and G/T were found in the six MeLiM pigs.

Tableau XXII: The relative frequency of the paralogous sequences of SNP 1535

Assay	Sample ID	Var. Pos.	A (%)	G (%)	~ A : G	# of A/A	# of A/G
SNP1535	B52-1535	R	64.32	35.68	2 : 1	1	2
	C284-1535	R	56.41	43.59	1 : 1	0	1
	C321-1535	R	61.79	38.21	3 : 2	1	4
	F206-1535	R	84.61	15.39	5 : 1	2	1
	485-1535	R	69.15	30.85	7 : 3	2	3
	486-1535	R	59.65	40.35	3 : 2	1	4
	LW5599-1535	R	100	0	1 : 0	1	0

Our results demonstrate that pyrosequencing could be successfully used and allowed us to validate SLA-7 CNVs in pigs. The results are in agreement with one copy for the reference haplotype Hp1a.0, as expected, and two or more copies in MeLiM animals. Additional experiments are required to further assess the exact number of SLA-7 paralogues in MeLiM pigs and we are aware that more validations are required.

Tableau XXIII: The relative frequency of the paralogous sequences of SNP1540

Assay	Sample ID	Var. Pos.	G (%)	T (%)	~ G : T	# of G/G	# of G/T
SNP1540	B52-1540	K	76.41	23.59	3 : 1	1	1
	C284-1540	K	70.38	29.62	7 : 3	2	3
	C321-1540	K	75.7	24.3	3 : 1	1	1
	F206-1540	K	88.25	11.75	9 : 1	4	1
	485-1540	K	77.3	22.7	3 : 1	1	1
	486-1540	K	72.91	27.09	7 : 3	2	3
	LW5599-1540	K	97.33	2.67	1 : 0	1	0

Discussion

There are two primary classes of MHC molecules, class I and class II. Among the class I molecules, the non classical class Ib are enigmatic and poorly studied in swine in contrast to the huge amount of work devoted to studies on human HLA-Ib molecules, especially HLA-E and HLA-G and to a lesser extent HLA-F. In pig, the gene structure, polymorphisms and function of SLA Ia molecules are well known and the three genes SLA-1, -2 and -3 stand for the functional homologue genes of the human HLA-A, -B and -C genes. However, despite these functional homologies, is noteworthy that the swine nomenclature does not follow the human nomenclature because no orthology has been established at the sequence level between each component of the two gene sets. In pig, the three genes SLA-6, -7 and -8 have been classified as non classical due to their genomic structure (Chardon et al. 2001) and limited polymorphism (MHD IPD database: <http://www.ebi.ac.uk/ipd/mhc/>). No orthology could be drawn between HLA-Ib and SLA-Ib genes (Chardon et al. 2001, Renard et al. 2006) and knowing if molecules with similar functions than HLA-Ib molecules in the pig is a main issue.

1- Polymorphism of SLA-Ib genes: characterization of SNPs and CNVs

Various types of polymorphisms exist in genomes, including SNPs, insertions, deletions, CNVs. The MHC genes are the most polymorphic genes known for vertebrates and the high number of polymorphisms in the MHC is often explained in terms of increased protection of hosts against pathogens (Borghans et al. 2004). Conversely, the limited polymorphism of

class Ib molecules is often related to the capacity to present a more limited variety of peptides than the class Ia molecules (Shawar et al. 1994). Our overall results have confirmed a limited polymorphism of SLA Ib genes that is in agreement with their classification as non-classical genes. Moreover, we have identified CNVs for SLA-7, suggesting a more complex polymorphic pattern that previously expected for this gene.

For MHC genes, cSNPs are mostly studied and exons 2 and 3 are known to be the most polymorphic by inducing amino acid changes that reshape the peptide binding groove. In human, the HLA Ia genes have a much higher polymorphism than HLA Ib genes. To date, the alleles of HLA-A, -B and -C were 1519, 2069 and 1016 respectively, but the alleles of HLA-E, -F, and -G were only 10, 22 and 46 (IMGT/HLA Database, <http://www.ebi.ac.uk/imgt/hla/>). In swine, fewer polymorphisms of MHC class Ia molecules than in human are reported but the allele numbers are higher than other genes, including SLA Ib genes. In our study, MeLiM pigs were genotyped for SLA genes for the first time. By comparing MeLiM pigs' SLA-7 sequences to the reference haplotype Hp1a.0, 14 SNPs were identified in exons, and 42 SNPs were found in the introns or the UTRs. The SNP polymorphism was not concentrated in exons 2 and 3 as expected for class Ia genes. For SLA-8, only 4 SNPs were detected in exons, and 56 SNPs were found in introns or regulatory sequences. Our results suggest that SLA-8 is less polymorphic than SLA-7. However, within exons 2 and 3 a total of six and three SNPs were detected for SLA-7 and -8, respectively. In case SLA-7 and -8 molecules present peptides to cell surface, the range of presented peptides might be more limited for SLA-8 than for SLA-7.

CNVs have been described for many types of genomic sequences and stand for DNA segments spanning several kilobases to megabases that vary in copy numbers among

individuals (Guryev et al. 2008, Orozco et al. 2009, Perry et al. 2008). These CNVs may include gene duplications and such structural variants are frequent in the pig genome (Ramayo-Caldas et al. 2010) as in other mammalian genomes (Guryev et al. 2008, Orozco et al. 2009, Perry et al. 2008). CNVs of SLA class I genes have already been reported using restriction fragment length polymorphisms (RFLP) analysis in different serologically defined haplotypes (Ruohonen-Lehto et al. 1998). Sequencing data from BAC clones corresponding to various SLA haplotypes have also revealed duplication of SLA-1 gene in some haplotypes (Tanaka-Matsuda et al. 2009). Our work on SNP characterization in SLA-Ib genes opened up new experiments with the aim to detect CNVs. Indeed, SLA-7 SNP distribution strongly suggested that more than one copy of the SLA-7 gene exist in MeLiM pigs.

To date, despite new technologies have been developed to determine the CNVs, such as tiling arrays (Fadista et al. 2008) or next generation sequencing, direct PCR-sequencing is still a simple, powerful and cost-effective method to study the CNV of a specific gene or a specific region of the genome. Therefore, we could use the data provided by long PCR and sequencing to characterize more than two alleles per animal that is in agreement with more than one SLA-7 gene. Moreover, the re-sequencing results of another animal (MeLiM #484) not included into the SNP analysis, also indicated several copies of SLA-7. A total of five haplotypes of SLA-7 could be characterized for SLA-7 in the MeLiM pig #484, suggesting three copies of SLA-7 gene at least. In order to confirm these CNV results obtained from direct sequencing data, we further used the pyrosequencing method known as an efficient SNP genotyping method for few SNPs and limited series of animals.

Pyrosequencing is a DNA sequencing technique based on the principle of sequencing-by-synthesis and detection of pyrophosphate release on nucleotide incorporation through a

series of enzymatic reactions (Doostzadeh et al. 2008). It is a real-time sequencing method (Ronaghi et al. 1998) that has been shown to be very effective in SNP dosage. At a SNP position, the percentage of each expected base provides the frequency of this base at the SNP position and therefore is suitable for confirmation/detection of duplications. Pyrosequencing has been extensively used in human single nucleotide polymorphism allele frequency studies (Doostzadeh et al. 2008, Gruber et al. 2002, Lavebratt and Sengul 2006). To our knowledge, this method was applied to the characterization of CNVs of MHC class I genes in pigs for the first time and we show that up to three copies of SLA-7 may exist in MeLiM pigs. Variations in SLA-Ia loci (SL A-1, -3, -5, and -9) have been reported in haplotypes Hp-28.0 and Hp-62.0 (Ruohonen-Lehto et al. 1998) and we demonstrate that variations in gene numbers also exist for non classical genes known to be functional as SLA-7. Variations in class Ib gene copies have been reported in cattle (Birch et al. 2006) and rat (Roos and Walter 2005).

Our results show that SLA-7 is duplicated in MeLiM pigs but that SLA-8 is likely not. We still miss data to know if SLA-6 may present CNVs. The study for CNV of SLA-6 and SLA-8 will be carried out in the future. We actually do not know if the duplicated copies are equally expressed and additional work is required to study allelic and locus-specific expression.

This study was an early step toward a more complete understanding of the polymorphism and copy number variation in pig MHC. Further studies on SLA genetic variations at the whole locus scale are being carried out in the laboratory. An ongoing program aims at specifically capturing the genomic segments corresponding to various MHC haplotypes for resequencing. A main issue will be to characterize SLA haplotypes by merging all relevant

existing polymorphisms of SLA-Ia and -Ib genes to give a more complete understanding of the MHC-related immune system and better contribute to improve health and resistance to pathogens in pig industry.

2- A splicing pattern with various complexities according to genes

Three different types of splicing patterns were identified for each of the three SLA-Ib genes. For SLA-8, a unique transcript of eight exons that encodes a protein of 355 aminoacids was identified. By contrast, five and six different RNA variants were characterized for SLA-6 and SLA-7, respectively. RNA splicing seems to be as a straightforward means to increase functional capacities of genes without increasing the gene number in genomes. Indeed, there seems to be an inverse proportion between the number of protein coding genes and the importance of splicing events as well as the raising credit of non coding RNA genes in the regulation of genome expression (Claverie 2005). Between one-third and two-thirds of human genes are estimated to produce at least two alternatively spliced isoforms (Brett et al. 2000, Croft et al. 2000, Kan et al. 2001, Lander et al. 2001, Mironov et al. 1999, Modrek et al. 2001) and 90,000 distinct alternative splicing events have been predicted for human genes (Kim et al. 2007). An overview of splicing gene patterns has been recently published for the pig genome and the occurrence of splicing events seems to follow the same general rules. At least 30% of the pig genes are expected to express alternative variants with tissue specificities (Nygard et al. 2010). It is noteworthy that the three SLA-Ib genes do not follow the same general rules for RNA transcription and it might be related to variable flexibility in their respective biological functions.

There are three transcripts of SLA-8 (SLA-8-001, OTTSUST00000000786; SLA-8-002, OTTSUST00000000787; and SLA-8-003, OTTSUST00000000788) in the VEGA database. Among them, SLA-8-001 is a protein-coding gene, and no proteins can be encoded by other two transcripts. In the MeLiM pig, only one transcript was found. And it has the same gene structure as the SLA-8-001 and can be translated to a 355 amino acid protein. This protein has a MHC class I complete structure, so it could bind a peptide to present the peptide to the cell surface with β -microglobulin.

For SLA-6, the RNA variants were clearly related either to exon skipping (SLA-6-1, -2 and -3) or to intron retention (SLA-6-4). The three variants SLA-6-2, -3 and -4 were found for the first time. SLA-6 has two transcripts (SLA-6-001, OTTSUST00000000780; SLA-6-002, OTTSUST00000000781) annotated in the VEGA database. In our study, SLA-6-001 stands for the variant SLA-6 and SLA-6-002 for the variant SLA-6-1 (see Figure XVI). Both transcripts are protein-coding genes, which contain complete open reading frames (ORF). These two transcripts were also identified in MeLiM pigs and the encoded proteins differ by the presence/absence of exon 2 that code for the alpha 1 domain of the class I molecules. The transcript SLA-6-2 lacks exons 2 and 6 and the transcript SLA-6-3 lacks exons 2, 5 and 6. Putative modified functions by comparison to the full length protein may be deduced from splicing events. SLA-6 isoforms may differ in the presence/absence of the alpha 2 domain that should result in modification of the ability to bind peptides. SLA-6 isoforms also differ in the presence/absence of a transmembrane domain that may lead to membrane bound or soluble proteins as reported for HLA-G molecules (Carosella et al. 2003). SLA-6 isoforms also differ in the length of the cytoplasmic tail reported to play a major regulatory role in extracellular expression by binding proteins in the cell cytoplasm (Gruda et al. 2007). The

structure of SLA-6-1, SLA-6-2 and SLA-6-3 makes the peptide presentation impossible. But if two SLA-6-1 molecules combine together or one SLA-6-1 combines with one SLA-6-2 or SLA-6-3 to form a homodimer or heterodimer, it would be possible to form a peptide groove and let the dimer have the ability of loading peptides as MHC class II (Busch et al. 2005, Ishitani and Geraghty 1992). Conversely to the variants SLA-6-1, -2 and -3, the variant SLA-6-4 is characterized by retention of exon 3 and a premature stop codon leads to a putative protein of 300 aminoacids. If translated, the biological meaning of this putative SLA-6 isoform is not easy to anticipate.

For SLA-7, four transcripts (SLA-7-001, OTTSUST00000000782; SLA-7-002, OTTSUST00000000783; SLA-7-003, OTTSUST00000000784; and SLA-7-004, OTTSUST00000000785) are reported in the VEGA database. SLA-7-001 and SLA-7-004 are protein-coding genes, but the protein encoded from SLA-7-004 only contains $\alpha 1$ and $\alpha 2$ domains. SLA-7-002 and SLA-7-003 are processed transcripts, which do not contain complete ORFs. In our study, six new SLA-7 transcripts were detected and characterized in the MeLiM pigs. The spliced variants are not simply explained by exon skipping or intron retention as for SLA-6. Firstly, a new transcript encoding a full length protein of eight exons has been identified and this result suggests that alternative SLA-7 full length protein isoforms exist. The last exon (exon 7) of SLA-7-001 (OTTSUST00000000782) was divided into two exons (exon 7 and exon 8) in MeLiM pig. The existence of an eighth exon for SLA-7 was identified for the first time in our study. Previously reported results together with ours suggest that two standard full length SLA-7 molecules may co-exist. We cannot rule out that these results are related to the possible duplications of SLA-7 genes in various SLA haplotypes as discussed earlier. We have also identified rarely used intron-exon-boundaries

for the spliced variant SLA-7-1 that encodes an SLA-6 isoform with a shortened alpha 3 domain (exon 4). The difference between the isoforms is that SLA-7-1 has a smaller $\alpha 3$ domain, with a length of 59 amino acids (39 amino acids encoded by exon 4a and 20 amino acids encoded by exon 4b). The $\alpha 3$ domain corresponds to the Immunoglobulin-like region and interacts with the cell surface of CD8 glycoproteins that are expressed on cytotoxic T lymphocytes and function as a co-receptor with the T cell receptor. The small $\alpha 3$ domain suggested that such a modification may alter interactions with cell receptors. This finding has been discussed in the manuscript that is in press (publication #1). The proteins encoded by the transcripts SLA-7-2 and SLA-7-3 cannot bind peptides for they only have an 18 amino acid leader peptide without a complete MHC molecular structure.

In conclusion, complexity of SLA-7 was much more complex than for SLA-8 and SLA-6. We cannot rule out that this complexity is related to SLA-7 duplications MeLiM pigs and that splicing events are gene copy-specific. More experiments are needed to address this question and it will be a major issue to know whether these findings are specific to MeLiM pigs or not.

3- Tissue specificity of SLA-Ib gene transcription and variations observed between MeLiM and Large White animals

We have initially studied tissue specificity of SLA-Ib transcription by designing primers that spanned the junction between exons 2 and 3 for each SLA-Ib gene (see results section and manuscript #2). Our results showed that the three genes are widely expressed in most tissues

with predominance in lung, lymphoid tissues and the digestive tract. Among the three genes, SLA-8 was found to be the most expressed. Our findings are in agreement with a less tissue restricted expression of SLA-Ib genes compared to HLA-Ib genes (Crew et al. 2004) and we could not find significant differences between MeLiM or Large White animals in this work. At the time these experiments were started, existence of complex splicing patterns of SLA-6 and -7 was not known and we should discuss more precisely now the respective transcription levels of the SLA-6, -7 and -8 genes in the tested tissues. The primers used for RT-qPCR could amplify the unique variant found for SLA-8 (see figure XVIII) and all SLA-7 RNA variants (see figure XVII) but only two variants (SLA-6 and SLA-6-4) among the five variants identified for SLA-6 (see figure XVI). We cannot rule out that some splice variants of SLA-6 that were not amplified with the primers used for RT-qPCR are highly expressed and these missing data could modify the respective order between the average transcription level of the three SLA-Ib genes that we have already reported. RT-PCRs specific for each SLA-6 splice variants have to be designed in order to address this question.

Most results on SLA splicing events have been reported using RT produced from total RNAs extracted from tissues of MeLiM animals. At first, we considered that most variations were expected in the gene segment spanning exons 4 to 8 and designed primers targeting this region. By using primers from exons 4 to the 3UTR, we found a splicing event after the stop that could not be due to a non sense mediated decay (NMD). The 650 nucleotide long sequence (SLA-7-650) corresponds to the expected full-length sequence. The 464 nucleotide long sequence SLA-7-464 (accession number: GU322919) is the result of an alternative splicing within the 3UTR region, 31 nucleotides downstream to the stop codon. Modification of the 3UTR does not affect the encoded molecules. The canonical GT-AG rule was used for

this splicing. A new category of transcripts has been recently characterized that can be subject to non sense mediated decay (NMD) (publication #1). Variants targeted by NMD can present alternative splicing in the 3'UTR and the distance between the stop codon and the splice site has been shown to be more than 50 nucleotides long (Mendell et al. 2004). The SLA-7-5 variant cannot fall into this category of transcripts subject to NMD because the distance between the stop codon and the splicing site is only 31 nucleotides. However, it is tempting to hypothesize that SLA-7-5 variants are subject to a post-transcriptional regulation and this phenomenon has to be explored. The expression patterns of these two 3UTR variants were studied in many different tissues including spleen, thymus, tonsil, and liver from MeLiM and Large White pigs (see manuscript 2 in press). Surprisingly, the 650 nt transcript band (existing in SLA-7, SLA-7-2, SLA-7-4) was detected in all tissues of both breeds but the SLA-7-5 was detected only in MeLiM pigs. We cannot rule out a very weak expression of the short variant in Large White pigs but our results strongly suggest a co-expression of both variants in MeLiM pigs and a predominant expression of the 650 nt transcript band in the Large White pigs included in our study. Our results show that transcripts with full-length 3'UTR (650 nt) and spliced UTR (464 nt) are co-expressed in most MeLiM tissues whereas the full-length 3UTR form is mostly found in large White pigs. As previously discussed, these findings may be due to SLA-7 locus specific expression. MeLiM pigs are minipigs that present a genetic susceptibility to spontaneously develop melanomas (Vincent-Naulleau et al. 2004) and regress in a constant time course manner that firstly involves genes related in cell cycle and secondly immunity-related genes such as SLA genes (Rambow et al. 2008a, Rambow et al. 2008b). We are planning to study the role of

non classical class I genes in the tumour progression/regression and include SLA-Ib genes as biologically relevant candidate genes.

4- Comparison of SLA Ib to HLA Ib genes

The MHC class Ib genes are intensively studied in human and it is highly tempting to analyze whether the class Ib genes in pig have similar functions and whether functional homologies can be hypothesized by looking gene per gene. No orthology could ever been identified between SLA-Ib and HLA-Ib genes neither at the gene position on the map nor at the sequence similarity (Chardon et al. 2001). In order to go further in possible functional homologies between pig and human genes, additional criteria can be used to propose some hypotheses that are not validated at the moment and are purely hypothetical. HLA-E has a main role in presenting leader peptides derived from HLA class I molecules (Hoare et al. 2006) and such a function could be conserved across many species. HLA-E has been shown to be the most transcribed in many tissues (Table II with EST counts in introduction) and the splicing pattern of this gene is limited (see Figure VII). HLA-G is highly specifically expressed and the splicing events lead to molecules that lack alpha 2 domain (exon 3), the transmembrane or modify the cytoplasmic tail (see Figure IX). HLA-F also presents a complex transcription pattern (see Figure VIII) and seems less specifically transcribed than HLA-G. Some HLA-H isoforms are soluble whereas such properties were not detected for HLA-E or HLA-F isoforms. All these data led us to the hypothesis that SLA-8, which is the most transcribed gene and presents the simplest transcription pattern could be a counterpart of HLA-E. Similarly, SLA-6 transcription pattern with putative soluble isoforms looks similar to HLA-G transcription pattern. These hypotheses are highly premature but are

tempting to propose. By referring to our results, SLA-8 would be the best candidate to look for expression on cell surface and see whether the HLA-E molecules can present peptides derived from the leader sequence of class I molecules.

5- Properties of the putative proteins encoded by SLA-Ib genes

The full-length transcripts including complete open reading frames were detected for SLA-6, -7 and -8 genes and the 3D-protein structures were predicted from the encoded proteins. They all contained extracellular $\alpha 1$, $\alpha 2$ and $\alpha 3$ domains, the transmembrane region and a cytoplasmic tail. The $\alpha 1$ and $\alpha 2$ domains compose the peptide-binding groove. This composition permits SLA Ib molecules to identify peptides and further present them to the cell surface. It seems that the three full length isoforms have the ability to bind to B2M and be expressed on cell surface. Therefore, the ability of peptide presentation is predicted with high probability. In order to address this question, we need specific antibodies and we are currently testing monoclonal antibodies that were produced by a private company. We have designed several cell expression systems and have to start again transient transfections of drosophila cells. The antibodies were produced by either a peptide or a cDNA strategy for immunization of mice. Sequences from MeLiM pigs were sent to the company with multi-alignments of SLA-Ia and -Ib genes. The antibodies expected to specifically recognize SLA-6 or SLA-8 molecules detect expression surface of molecules on the porcine PK15 cells by FACS but since the non transfected cells provide similar levels than the transfected cells, we suspect cross reactions with SLA-Ia molecules and cannot yet prove that SLA-6 or -8 are membrane bound glycoproteins. Experiments are in progress to transfect HeLa cells and the

drosophila cells in order to detect the protein expression in cellular context harbouring no background of SLA-Ia or -Ib expression. For the drosophila cells, we plan to co-transfect a vector expressing B2M with a vector expressing each SLA-Ib gene. Co-transfection of vectors expressing B2M and SLA-1 will stand for positive controls for membrane expression. These experiments are very promising and even if the monoclonal antibodies are not strictly molecule specific, we could go for functional studies using heterologous systems like SC2 or human cells more appropriated than HeLa cells.

As previously discussed, SLA-6 can encode isoforms that could be soluble as some HLA-G isoforms, due to the absence of a transmembrane domain (exon 5). Other isoforms present variations in the cytoplasmic tail. A perspective would be to construct expression vectors for each RNA variant and look for protein expression and properties.

Conclusion

Our work provides a wide range of new data on SLA-Ib genes that correspond to an original contribution to the field. A first set of key findings relates to the deep characterization of the splicing transcription patterns of the three SLA-Ib genes with various complexities according to genes. A second set of key findings specifically relates to the SLA-7 gene that could be re-annotated (8 exons instead of seven) and that was shown to harbour the highest complexity at the RNA and DNA level with an unexpected variation in the copy number and a discrepancy for expression of some transcripts between Large White and MeLiM pigs. A third set of key findings relates to the production of monoclonal antibodies specific for SLA-Ib molecules and our results may be considered as promising even if not fully completed yet.

We have mostly worked on adult tissues and the next steps will be to study the expression of SLA-Ib genes during embryonic development and at the foeto-maternal interface. In addition, the melanoma bearing MeLiM pigs provide a major biomedical model to study the expression of SLA-Ib genes in progressing and regressing tumours as well as in normal melanocytes and skin. Addressing the specific role of each SLA-Ib gene for presenting peptides and interacting with NK cells is a major question to address and our work is a first step towards such studies.

References

- Amills M, Ramiya V, Norimine J, Lewin HA. 1998. The major histocompatibility complex of ruminants. *Rev Sci Tech* 17: 108-120.
- Andersson G, Svensson AC, Setterblad N, Rask L. 1998. Retroelements in the human MHC class II region. *Trends Genet* 14: 109-114.
- Anfossi N, et al. 2006. Human NK cell education by inhibitory receptors for MHC class I. *Immunity* 25: 331-342.
- Anzai T, et al. 2003. Comparative sequencing of human and chimpanzee MHC class I regions unveils insertions/deletions as the major path to genomic divergence. *Proc Natl Acad Sci U S A* 100: 7708-7713.
- Apps R, Gardner L, Traherne J, Male V, Moffett A. 2008. Natural-killer cell ligands at the maternal-fetal interface: UL-16 binding proteins, MHC class-I chain related molecules, HLA-F and CD48. *Hum Reprod* 23: 2535-2548.
- Archbold JK, Ely LK, Kjer-Nielsen L, Burrows SR, Rossjohn J, McCluskey J, Macdonald WA. 2008. T cell allorecognition and MHC restriction--A case of Jekyll and Hyde? *Mol Immunol* 45: 583-598.
- Auchincloss H, Jr., Lee R, Shea S, Markowitz JS, Grusby MJ, Glimcher LH. 1993. The role of "indirect" recognition in initiating rejection of skin grafts from major histocompatibility complex class II-deficient mice. *Proc Natl Acad Sci U S A* 90: 3373-3377.
- Badley JE, Frelinger JA. 1991. Expression of a MHC non-classical class I gene, Q4, is similar to a classical class I gene, Dp. *Eur J Immunogenet* 18: 315-321.
- Ballingall KT, Ellis SA, MacHugh ND, Archibald SD, McKeever DJ. 2004a. The DY genes of the cattle MHC: expression and comparative analysis of an unusual class II MHC gene pair. *Immunogenetics* 55: 748-755.
- Ballingall KT, Luyai A, Rowlands GJ, Sales J, Musoke AJ, Morzaria SP, McKeever DJ. 2004b. Bovine leukocyte antigen major histocompatibility complex class II DRB3*2703 and DRB3*1501 alleles are associated with variation in levels of protection against *Theileria parva* challenge following immunization with the sporozoite p67 antigen. *Infect Immun* 72: 2738-2741.
- Band M, Larson JH, Womack JE, Lewin HA. 1998. A radiation hybrid map of BTA23: identification of a chromosomal rearrangement leading to separation of the cattle MHC class II subregions. *Genomics* 53: 269-275.
- Benacerraf B, Mcdevitt HO. 1972. Histocompatibility-Linked Immune-Response Genes. *Science* 175: 273.

Bendixen E, Danielsen M, Larsen K, Bendixen C. 2010. Advances in porcine genomics and proteomics--a toolbox for developing the pig as a model organism for molecular biomedical research. *Brief Funct Genomics* 9: 208-219.

Bijen CB, Bantema-Joppe EJ, de Jong RA, Leffers N, Mourits MJ, Eggink HF, van der Zee AG, Hollema H, de Bock GH, Nijman HW. 2010. The prognostic role of classical and nonclassical MHC class I expression in endometrial cancer. *Int J Cancer* 126: 1417-1427.

Birch J, Murphy L, MacHugh ND, Ellis SA. 2006. Generation and maintenance of diversity in the cattle MHC class I region. *Immunogenetics* 58: 670-679.

Birch J, Codner G, Guzman E, Ellis SA. 2008. Genomic location and characterisation of nonclassical MHC class I genes in cattle. *Immunogenetics* 60: 267-273.

Borghans JA, Beltman JB, De Boer RJ. 2004. MHC polymorphism under host-pathogen coevolution. *Immunogenetics* 55: 732-739.

Bos DH, Waldman B. 2006. Polymorphism, natural selection, and structural modeling of class Ia MHC in the African clawed frog (*Xenopus laevis*). *Immunogenetics* 58: 433-442.

Boyle LH, Gillingham AK, Munro S, Trowsdale J. 2006. Selective export of HLA-F by its cytoplasmic tail. *J Immunol* 176: 6464-6472.

Bradley KJ, Cavaco BM, Bowl MR, Harding B, Young A, Thakker RV. 2005. Utilisation of a cryptic non-canonical donor splice site of the gene encoding PARAFIBROMIN is associated with familial isolated primary hyperparathyroidism. *J Med Genet* 42: e51.

Brett D, Hanke J, Lehmann G, Haase S, Delbruck S, Krueger S, Reich J, Bork P. 2000. EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. *FEBS Lett* 474: 83-86.

Brinkmeyer-Langford C, Raudsepp T, Gustafson-Seabury A, Chowdhary BP. 2008. A BAC contig map over the proximal approximately 3.3 Mb region of horse chromosome 21. *Cytogenet Genome Res* 120: 164-172.

Brinkmeyer-Langford CL, Childers CP, Fritz KL, Gustafson-Seabury AL, Cothran M, Raudsepp T, Womack JE, Skow LC. 2009. A high resolution RH map of the bovine major histocompatibility complex. *BMC Genomics* 10: 182.

Busch R, Rinderknecht CH, Roh S, Lee AW, Harding JJ, Burster T, Hornell TM, Mellins ED. 2005. Achieving stability through editing and chaperoning: regulation of MHC class II peptide binding and expression. *Immunol Rev* 207: 242-260.

Carosella ED, Favier B, Rouas-Freiss N, Moreau P, Lemaoult J. 2008. Beyond the increasing complexity of the immunomodulatory HLA-G molecule. *Blood* 111: 4862-4870.

Carosella ED, Moreau P, Le Maoult J, Le Discorde M, Dausset J, Rouas-Freiss N. 2003. HLA-G molecules: from maternal-fetal tolerance to tissue acceptance. *Adv Immunol* 81: 199-252.

Chardon P, Rogel-Gaillard C, Cattolico L, Duprat S, Vaiman M, Renard C. 2001. Sequence of the swine major histocompatibility complex region containing all non-classical class I genes. *Tissue Antigens* 57: 55-65.

Claverie JM. 2005. Fewer genes, more noncoding RNA. *Science* 309: 1529-1530.

Collins T, Lapierre LA, Fiers W, Strominger JL, Pober JS. 1986. Recombinant human tumor necrosis factor increases mRNA levels and surface expression of HLA-A,B antigens in vascular endothelial cells and dermal fibroblasts in vitro. *Proc Natl Acad Sci U S A* 83: 446-450.

Comiskey M, Goldstein CY, De Fazio SR, Mammolenti M, Newmark JA, Warner CM. 2003. Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product. *Hum Immunol* 64: 999-1004.

Crew MD, Phanavanh B, Garcia-Borges CN. 2004. Sequence and mRNA expression of nonclassical SLA class I genes SLA-7 and SLA-8. *Immunogenetics* 56: 111-114.

Croft L, Schandorff S, Clark F, Burrage K, Arctander P, Mattick JS. 2000. ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome. *Nat Genet* 24: 340-341.

Danchin E, Vitiello V, Vienne A, Richard O, Gouret P, McDermott MF, Pontarotti P. 2004. The major histocompatibility complex origin. *Immunol Rev* 198: 216-232.

Dausset J. 1958. [Demonstration of the Rh (D) antigen on leukocytes and platelets from humans]. *C R Hebd Seances Acad Sci* 246: 2171-2173.

Davies B, Hiby S, Gardner L, Loke YW, King A. 2001. HLA-G expression by tumors. *Am J Reprod Immunol* 45: 103-107.

Derre L, Corvaisier M, Charreau B, Moreau A, Godefroy E, Moreau-Aubry A, Jotereau F, Gervois N. 2006. Expression and release of HLA-E by melanoma cells and melanocytes: potential impact on the response of cytotoxic effector cells. *J Immunol* 177: 3100-3107.

Dick TP, Bangia N, Peaper DR, Cresswell P. 2002. Disulfide bond isomerization and the assembly of MHC class I-peptide complexes. *Immunity* 16: 87-98.

Doostzadeh J, Shokralla S, Absalan F, Jalili R, Mohandessi S, Langston JW, Davis RW, Ronaghi M, Gharizadeh B. 2008. High throughput automated allele frequency estimation by pyrosequencing. *PLoS One* 3: e2693.

Dukkipati VS, Blair HT, Garrick DJ, Murray A. 2006. 'Ovar-Mhc' - ovine major histocompatibility complex: structure and gene polymorphisms. *Genet Mol Res* 5: 581-608.

Ehrlich R, Lifshitz R, Pescovitz MD, Rudikoff S, Singer DS. 1987. Tissue-specific expression and structure of a divergent member of a class I MHC gene family. *J Immunol* 139: 593-602.

- Fadista J, Nygaard M, Holm LE, Thomsen B, Bendixen C. 2008. A snapshot of CNVs in the pig genome. *PLoS One* 3: e3916.
- Gao J, Liu K, Liu H, Blair HT, Li G, Chen C, Tan P, Ma RZ. 2010. A complete DNA sequence map of the ovine major histocompatibility complex. *BMC Genomics* 11: 466.
- Garbi N, Tan P, Diehl AD, Chambers BJ, Ljunggren HG, Momburg F, Hammerling GJ. 2000. Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. *Nat Immunol* 1: 234-238.
- Garrido F, Ruiz-Cabello F, Cabrera T, Perez-Villar JJ, Lopez-Botet M, Duggan-Keen M, Stern PL. 1997. Implications for immunosurveillance of altered HLA class I phenotypes in human tumours. *Immunol Today* 18: 89-95.
- Geffrotin C, et al. 2004. Identification of five chromosomal regions involved in predisposition to melanoma by genome-wide scan in the MeLiM swine model. *Int J Cancer* 110: 39-50.
- Germain RN, Margulies DH. 1993. The biochemistry and cell biology of antigen processing and presentation. *Annu Rev Immunol* 11: 403-450.
- Gould DS, Auchincloss H, Jr. 1999. Direct and indirect recognition: the role of MHC antigens in graft rejection. *Immunol Today* 20: 77-82.
- Gruber JD, Colligan PB, Wolford JK. 2002. Estimation of single nucleotide polymorphism allele frequency in DNA pools by using Pyrosequencing. *Hum Genet* 110: 395-401.
- Gruda R, Achdout H, Stern-Ginossar N, Gazit R, Betser-Cohen G, Manaster I, Katz G, Gonen-Gross T, Tirosh B, Mandelboim O. 2007. Intracellular cysteine residues in the tail of MHC class I proteins are crucial for extracellular recognition by leukocyte Ig-like receptor 1. *J Immunol* 179: 3655-3661.
- Guleria I, Sayegh MH. 2007. Maternal acceptance of the fetus: true human tolerance. *J Immunol* 178: 3345-3351.
- Guryev V, et al. 2008. Distribution and functional impact of DNA copy number variation in the rat. *Nat Genet* 40: 538-545.
- Gustafson AL, Tallmadge RL, Ramlachan N, Miller D, Bird H, Antczak DF, Raudsepp T, Chowdhary BP, Skow LC. 2003. An ordered BAC contig map of the equine major histocompatibility complex. *Cytogenet Genome Res* 102: 189-195.
- Haeney M. 1995. The immunological background to transplantation. *J Antimicrob Chemother* 36 Suppl B: 1-9.
- Ho CS, Lunney JK, Ando A, Rogel-Gaillard C, Lee JH, Schook LB, Smith DM. 2009. Nomenclature for factors of the SLA system, update 2008. *Tissue Antigens* 73: 307-315.
- Hoare HL, et al. 2006. Structural basis for a major histocompatibility complex class Ib-restricted T cell response. *Nat Immunol* 7: 256-264.

- Horak V, Fortyn K, Hruban V, Klaudy J. 1999. Hereditary melanoblastoma in miniature pigs and its successful therapy by devitalization technique. *Cell Mol Biol (Noisy-le-grand)* 45: 1119-1129.
- Horton R, et al. 2004. Gene map of the extended human MHC. *Nat Rev Genet* 5: 889-899.
- Hunt JS, Morales PJ, Pace JL, Fazleabas AT, Langat DK. 2007. A commentary on gestational programming and functions of HLA-G in pregnancy. *Placenta* 28 Suppl A: S57-63.
- Hurt P, et al. 2004. The genomic sequence and comparative analysis of the rat major histocompatibility complex. *Genome Res* 14: 631-639.
- Ishitani A, Geraghty DE. 1992. Alternative splicing of HLA-G transcripts yields proteins with primary structures resembling both class I and class II antigens. *Proc Natl Acad Sci U S A* 89: 3947-3951.
- Ishitani A, Sageshima N, Lee N, Dorofeeva N, Hatake K, Marquardt H, Geraghty DE. 2003. Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. *J Immunol* 171: 1376-1384.
- Ito T, Ito N, Saathoff M, Stampachiachiere B, Bettermann A, Bulfone-Paus S, Takigawa M, Nickoloff BJ, Paus R. 2005. Immunology of the human nail apparatus: the nail matrix is a site of relative immune privilege. *J Invest Dermatol* 125: 1139-1148.
- Ivanyi J, Pinter M. 1967. [The splenogram in acute viral hepatitis]. *Acta Hepatosplenol* 14: 237-241.
- Ivanyi P. 1970. The major histocompatibility antigens in various species. *Curr Top Microbiol Immunol* 53: 1-90.
- Jarrell VL, Lewin HA, Da Y, Wheeler MB. 1995. Gene-centromere mapping of bovine DYA, DRB3, and PRL using secondary oocytes and first polar bodies: evidence for four-strand double crossovers between DYA and DRB3. *Genomics* 27: 33-39.
- Joly E, Rouillon V. 2006. The orthology of HLA-E and H2-Qa1 is hidden by their concerted evolution with other MHC class I molecules. *Biol Direct* 1: 2.
- Kan Z, Rouchka EC, Gish WR, States DJ. 2001. Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res* 11: 889-900.
- Kelley J, Walter L, Trowsdale J. 2005. Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56: 683-695.
- Kevorkian S. 2007. Etude de la transcription des gènes d'histocompatibilité de class I non classiques SLA-6, SLA-7 et SLA-8 du porc. Master thesis.
- Kim N, Alekseyenko AV, Roy M, Lee C. 2007. The ASAP II database: analysis and comparative genomics of alternative splicing in 15 animal species. *Nucleic Acids Res* 35: D93-98.

Klein J. 1986. Seeds of time: fifty years ago Peter A. Gorer discovered the H-2 complex. *Immunogenetics* 24: 331-338.

Koller BH, Geraghty DE, Shimizu Y, DeMars R, Orr HT. 1988. HLA-E. A novel HLA class I gene expressed in resting T lymphocytes. *J Immunol* 141: 897-904.

Lamm LU, Olaisen B. 1985. Report of the Committee on the Genetic Constitution of Chromosomes 5 and 6. *Cytogenet Cell Genet* 40: 128-155.

Lander ES, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.

Lanier LL. 1998. NK cell receptors. *Annu Rev Immunol* 16: 359-393.

Lankat-Buttgereit B, Tampe R. 2002. The transporter associated with antigen processing: function and implications in human diseases. *Physiol Rev* 82: 187-204.

Lau P, Amadou C, Brun H, Rouillon V, McLaren F, Le Rolle AF, Graham M, Butcher GW, Joly E. 2003. Characterisation of RT1-E2, a multigenic family of highly conserved rat non-classical MHC class I molecules initially identified in cells from immunoprivileged sites. *BMC Immunol* 4: 7.

Lavebratt C, Sengul S. 2006. Single nucleotide polymorphism (SNP) allele frequency estimation in DNA pools using Pyrosequencing. *Nat Protoc* 1: 2573-2582.

Le Bouteiller P. 1994. HLA class I chromosomal region, genes, and products: facts and questions. *Crit Rev Immunol* 14: 89-129.

Le Bouteiller P, Pizzato N, Barakonyi A, Solier C. 2003. HLA-G, pre-eclampsia, immunity and vascular events. *J Reprod Immunol* 59: 219-234.

Le Discorde M, Moreau P, Sabatier P, Legeais JM, Carosella ED. 2003. Expression of HLA-G in human cornea, an immune-privileged tissue. *Hum Immunol* 64: 1039-1044.

Le Maoult J, Rouas-Freiss N, Le Discorde M, Moreau P, Carosella ED. 2004. [HLA-G in organ transplantation]. *Pathol Biol (Paris)* 52: 97-103.

Lee C, Liang MN, Tate KM, Rabinowitz JD, Beeson C, Jones PP, McConnell HM. 1998. Evidence that the autoimmune antigen myelin basic protein (MBP) Ac1-9 binds towards one end of the major histocompatibility complex (MHC) cleft. *Journal of Experimental Medicine* 187: 1505-1516.

Lewin HA, Russell GC, Glass EJ. 1999. Comparative organization and function of the major histocompatibility complex of domesticated cattle. *Immunol Rev* 167: 145-158.

Long EO. 1999. Regulation of immune responses through inhibitory receptors. *Annual Review of Immunology* 17: 875-904.

Lunney JK, Ho CS, Wysocki M, Smith DM. 2009. Molecular genetics of the swine major histocompatibility complex, the SLA complex. *Dev Comp Immunol* 33: 362-374.

- Margalit Y, Yarus S, Shapira E, Gruenbaum Y, Fainsod A. 1993. Isolation and characterization of target sequences of the chicken CdxA homeobox gene. *Nucleic Acids Res* 21: 4915-4922.
- Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 36: 1073-1078.
- MHC-Sequencing-Consortium. 1999. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* 401: 921-923.
- Miller DM, Rahill BM, Boss JM, Lairmore MD, Durbin JE, Waldman JW, Sedmak DD. 1998. Human cytomegalovirus inhibits major histocompatibility complex class II expression by disruption of the Jak/Stat pathway. *Journal of Experimental Medicine* 187: 675-683.
- Mironov AA, Fickett JW, Gelfand MS. 1999. Frequent alternative splicing of human genes. *Genome Res* 9: 1288-1293.
- Mizuno S, Trapani JA, Koller BH, Dupont B, Yang SY. 1988. Isolation and nucleotide sequence of a cDNA clone encoding a novel HLA class I gene. *J Immunol* 140: 4024-4030.
- Modrek B, Resch A, Grasso C, Lee C. 2001. Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res* 29: 2850-2859.
- Moretta A, Vitale M, Sivori S, Morelli L, Pende D, Bottino C. 1996a. Inhibitory and activatory receptors for HLA class I molecules in human natural killer cells. *Chem Immunol* 64: 77-87.
- Moretta A, Bottino C, Vitale M, Pende D, Biassoni R, Mingari MC, Moretta L. 1996b. Receptors for HLA class-I molecules in human natural killer cells. *Annu Rev Immunol* 14: 619-648.
- Moretta L, Moretta A. 2004. Unravelling natural killer cell function: triggering and inhibitory human NK receptors. *EMBO J* 23: 255-259.
- Moretta L, Bottino C, Pende D, Vitale M, Mingari MC, Moretta A. 2004. Different checkpoints in human NK-cell activation. *Trends Immunol* 25: 670-676.
- Moscoso J, Serrano-Vela JI, Pacheco R, Arnaiz-Villena A. 2006a. HLA-G, -E and -F: allelism, function and evolution. *Transpl Immunol* 17: 61-64.
- Moscoso J, Seclen S, Serrano-Vela JI, Villena A, Martinez-Laso J, Zamora J, Moreno A, Ira-Cachafeiro J, Arnaiz-Villena A. 2006b. HLA genes in Lamas Peruvian-Amazonian Amerindians. *Mol Immunol* 43: 1881-1889.
- Nygard AB, Cirera S, Gilchrist MJ, Gorodkin J, Jorgensen CB, Fredholm M. 2010. A study of alternative splicing in the pig. *BMC Res Notes* 3: 123.

O'Callaghan CA, Tormo J, Willcox BE, Braud VM, Jakobsen BK, Stuart DI, McMichael AJ, Bell JI, Jones EY. 1998. Structural features impose tight peptide binding specificity in the nonclassical MHC molecule HLA-E. *Mol Cell* 1: 531-541.

Orozco LD, Cokus SJ, Ghazalpour A, Ingram-Drake L, Wang S, van Nas A, Che N, Araujo JA, Pellegrini M, Lusis AJ. 2009. Copy number variation influences gene expression and metabolic traits in mice. *Hum Mol Genet* 18: 4118-4129.

Pedersen MO, Underhaug J, Dittmer J, Miller M, Nielsen NC. 2008. The three-dimensional structure of CsmA: a small antenna protein from the green sulfur bacterium *Chlorobium tepidum*. *FEBS Lett* 582: 2869-2874.

Perry GH, et al. 2008. The fine-scale and complex architecture of human copy-number variation. *Am J Hum Genet* 82: 685-695.

Ponte M, Cantoni C, Biassoni R, Tradori-Cappai A, Bentivoglio G, Vitale C, Bertone S, Moretta A, Moretta L, Mingari MC. 1999. Inhibitory receptors sensing HLA-G1 molecules in pregnancy: decidua-associated natural killer cells express LIR-1 and CD94/NKG2A and acquire p49, an HLA-G1-specific receptor. *Proc Natl Acad Sci U S A* 96: 5674-5679.

Rabin M, Fries R, Singer D, Ruddle FH. 1985. Assignment of the porcine major histocompatibility complex to chromosome 7 by in situ hybridization. *Cytogenet Cell Genet* 39: 206-209.

Radcliffe CM, Diedrich G, Harvey DJ, Dwek RA, Cresswell P, Rudd PM. 2002. Identification of specific glycoforms of major histocompatibility complex class I heavy chains suggests that class I peptide loading is an adaptation of the quality control pathway involving calreticulin and ERp57. *J Biol Chem* 277: 46415-46423.

Raffaghello L, et al. 2007. Expression and functional analysis of human leukocyte antigen class I antigen-processing machinery in medulloblastoma. *Cancer Res* 67: 5471-5478.

Rajagopalan S, Long EO. 1999. A human histocompatibility leukocyte antigen (HLA)-G-specific receptor expressed on all natural killer cells. *Journal of Experimental Medicine* 189: 1093-1100.

Ramayo-Caldas Y, Castello A, Pena RN, Alves E, Mercade A, Souza CA, Fernandez AI, Perez-Enciso M, Folch JM. 2010. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* 11: 593.

Rambow F, Piton G, Bouet S, Leplat JJ, Baulande S, Marrau A, Stam M, Horak V, Vincent-Naulleau S. 2008a. Gene expression signature for spontaneous cancer regression in melanoma pigs. *Neoplasia* 10: 714-726, 711 p following 726.

Rambow F, Malek O, Geffrotin C, Leplat JJ, Bouet S, Piton G, Hugot K, Bevilacqua C, Horak V, Vincent-Naulleau S. 2008b. Identification of differentially expressed genes in spontaneously regressing melanoma using the MeLiM swine model. *Pigment Cell Melanoma Res* 21: 147-161.

Renard C, et al. 2006. The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics* 88: 96-110.

Ritz U, Momburg F, Pilch H, Huber C, Maeurer MJ, Seliger B. 2001. Deficient expression of components of the MHC class I antigen processing machinery in human cervical carcinoma. *Int J Oncol* 19: 1211-1220.

Rodgers JR, Cook RG. 2005. MHC class Ib molecules bridge innate and acquired immunity. *Nat Rev Immunol* 5: 459-471.

Rolland P, Deen S, Scott I, Durrant L, Spendlove I. 2007. Human leukocyte antigen class I antigen expression is an independent prognostic factor in ovarian cancer. *Clin Cancer Res* 13: 3591-3596.

Ronaghi M, Uhlen M, Nyren P. 1998. A sequencing method based on real-time pyrophosphate. *Science* 281: 363, 365.

Roos C, Walter L. 2005. Considerable haplotypic diversity in the RT1-CE class I gene region of the rat major histocompatibility complex. *Immunogenetics* 56: 773-777.

Rouas-Freiss N, Moreau P, Ferrone S, Carosella ED. 2005. HLA-G proteins in cancer: do they provide tumor cells with an escape mechanism? *Cancer Res* 65: 10139-10144.

Ruohonen-Lehto MK, Renard C, Rothschild MF, Edfors-Lilja I, Kristensen B, Gustafsson U, Larson RG, Varvio SL. 1998. Variable number of pig MHC class I genes in different serologically defined haplotypes identified by a 3'-untranslated region probe. *Anim Genet* 29: 178-184.

Salter-Cid L, Nonaka M, Flajnik MF. 1998. Expression of MHC class Ia and class Ib during ontogeny: high expression in epithelia and coregulation of class Ia and *Imp7* genes. *J Immunol* 160: 2853-2861.

Saric T, Chang SC, Hattori A, York IA, Markant S, Rock KL, Tsujimoto M, Goldberg AL. 2002. An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat Immunol* 3: 1169-1176.

Schutze H, Gorer PA, Finlayson MH. 1936. The Resistance of Four Mouse Lines to Bacterial Infection. *J Hyg (Lond)* 36: 37-49.

Seliger B, Abken H, Ferrone S. 2003. HLA-G and MIC expression in tumors and their role in anti-tumor immunity. *Trends Immunol* 24: 82-87.

Serwold T, Gonzalez F, Kim J, Jacob R, Shastri N. 2002. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* 419: 480-483.

Shawar SM, Vyas JM, Rodgers JR, Rich RR. 1994. Antigen presentation by major histocompatibility complex class I-B molecules. *Annu Rev Immunol* 12: 839-880.

Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR, Cerrina F. 1999. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat Biotechnol* 17: 974-978.

Skow LC, Snaples SN, Davis SK, Taylor JF, Huang B, Gallagher DH. 1996. Localization of bovine lymphocyte antigen (BoLA) DYA and class I loci to different regions of chromosome 23. *Mamm Genome* 7: 388-389.

Snell GD. 1976. Recognition structures determined by the H-2 complex. *Transplant Proc* 8: 147-156.

Sullivan CA, Ghosh S, Ocal IT, Camp RL, Rimm DL, Chung GG. 2009. Microvessel area using automated image analysis is reproducible and is associated with prognosis in breast cancer. *Hum Pathol* 40: 156-165.

Sullivan LC, Clements CS, Rossjohn J, Brooks AG. 2008. The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. *Tissue Antigens*.

Tanaka-Matsuda M, Ando A, Rogel-Gaillard C, Chardon P, Uenishi H. 2009. Difference in number of loci of swine leukocyte antigen classical class I genes among haplotypes. *Genomics* 93: 261-273.

Tennant LM, Renard C, Chardon P, Powell PP. 2007. Regulation of porcine classical and nonclassical MHC class I expression. *Immunogenetics* 59: 377-389.

The MHC sequencing consortium. 1999. Complete sequence and gene map of a human major histocompatibility complex. *Nature* 401: 921-923.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22(22):4673-80.

Trgovcich J, Cebulla C, Zimmerman P, Sedmak DD. 2006. Human cytomegalovirus protein pp71 disrupts major histocompatibility complex class I cell surface expression. *J Virol* 80: 951-963.

Trowsdale J, Moffett A. 2008. NK receptor interactions with MHC class I molecules in pregnancy. *Semin Immunol* 20: 317-320.

Uhrberg M, Valiante NM, Young NT, Lanier LL, Phillips JH, Parham P. 2001. The repertoire of killer cell Ig-like receptor and CD94:NKG2A receptors in T cells: clones sharing identical alpha beta TCR rearrangement express highly diverse killer cell Ig-like receptor patterns. *J Immunol* 166: 3923-3932.

Vaiman M, Renard C, LaFage P, Ameteau J, Nizza P. 1970. Evidence for a histocompatibility system in swine (SL-A). *Transplantation* 10: 155-164.

VEGA_database. The Vertebrate Genome Annotation (VEGA) database. <http://vega.sanger.ac.uk/index.html>.

- Vincent-Naulleau S, et al. 2004. Clinical and histopathological characterization of cutaneous melanomas in the melanoblastoma-bearing Libechev minipig model. *Pigment Cell Res* 17: 24-35.
- Wagner JL. 2003. Molecular organization of the canine major histocompatibility complex. *J Hered* 94: 23-26.
- Wainwright SD, Biro PA, Holmes CH. 2000. HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. *J Immunol* 164: 319-328.
- Waterston RH, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-562.
- Wright H, Ballingall KT. 1994. Mapping and characterization of the DQ subregion of the ovine MHC. *Anim Genet* 25: 243-249.
- Wright H, Ballingall KT, Redmond J. 1994. The DY sub-region of the sheep MHC contains an A/B gene pair. *Immunogenetics* 40: 230-234.
- Yewdell JW, Bennink JR. 2001. Cut and trim: generating MHC class I peptide ligands. *Curr Opin Immunol* 13: 13-18.
- Yewdell JW, Reits E, Neefjes J. 2003. Making sense of mass destruction: quantitating MHC class I antigen presentation. *Nat Rev Immunol* 3: 952-961.
- Zinkernagel RM, Doherty PC. 1974. Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. *Nature* 251: 547-548.

Publication # 1

Accepted for publication in BMC Proceedings

Transcription variants of *SLA-7*, a swine non classical MHC class I gene

Rui Hu, Gaëtan Lemonnier, Emmanuelle Bourneuf, Silvia Vincent-Naulleau, Claire Rogel-Gaillard

Transcription variants of *SLA-7*, a swine non classical MHC class I gene

Rui Hu^{1,2,3}, Gaëtan Lemonnier^{1,2,3}, Emmanuelle Bourneuf^{1,2,3}, Silvia Vincent-Naulleau^{1,2,3},
Claire Rogel-Gaillard^{1,2,3§}

Accepted for publication in BMC Proceedings

¹INRA, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

²CEA, DSV, iRCM, Laboratoire de Radiobiologie et Etude du Génome, Jouy-en-Josas,
France

³AgroParisTech, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

§ Corresponding author

Email addresses:

RH: rui.hu@jouy.inra.fr

GL: gaetan.lemonnier@jouy.inra.fr

EB: emmanuelle.bourneuf@cea.fr

SVN: silvia.vincentn@cea.fr

CRG: claire.rogel-gaillard@jouy.inra.fr

Abstract

In pig, very little information is available on the non classical class I (Ib) genes of the Major Histocompatibility Complex (MHC) i.e. *SLA-6*, *-7* and *-8*. Our aim was to focus on the transcription pattern of the *SLA-7* gene. RT-PCR experiments were carried out with *SLA-7* specific primers targeting either the full coding sequence (CDS) from exon 1 to the 3 prime untranslated region (3UTR) or a partial CDS from exon 4 to the 3UTR. We show that the *SLA-7* gene expresses a full length transcript not yet identified that refines annotation of the gene with eight exons instead of seven as initially described from the existing RefSeq RNA. These two RNAs encode molecules that differ in cytoplasmic tail length. In this study, another *SLA-7* transcript variant was characterized, which encodes a protein with a shorter alpha 3 domain, as a consequence of a splicing site within exon 4. Surprisingly, a cryptic non canonical GA-AG splicing site is used to generate this transcript variant. An additional *SLA-7* variant was also identified in the 3UTR with a splicing site occurring 31 nucleotides downstream to the stop codon. In conclusion, the pig MHC class Ib gene presents a complex transcription pattern with two transcripts encoding various molecules and transcripts that do not alter the CDS and may be subject to post-transcriptional regulation.

Background

The Major Histocompatibility Complex (MHC) class I gene family comprises classical (Ia) and non classical (Ib) genes. The highly polymorphic class Ia genes are widely expressed and encode membrane-bound glycoproteins that present self and viral peptides to cytotoxic T cells [1] and modulate the activity of natural killer cells [2]. In contrast, the class Ib genes display limited polymorphism, and are predominantly expressed in immunotolerant organ sites in human, notably at the feto-maternal interface [3]. In man, three MHC class Ib genes have been characterized, namely *HLA-E*, *-F* and *-G* [3] and *HLA-G* has been shown to express alternatively spliced variants encoding various membrane-bound as well as soluble proteins [4]. In mouse, the *H2-Qa1* gene is orthologous to *HLA-E* [5] and functional homologies have been established between *H2-Qa2* and *HLA-G* [6]. One to four MHC class Ib genes have been identified in rat according to haplotypes [7] and four MHC class Ib genes have been characterized in cattle [8]. There is a growing interest in addressing the role of the MHC class Ib genes in the species where they are characterized. Indeed, MHC class Ia genes seem to share similar functions across species but the MHC class Ib genes are good candidates to address questions on both shared and species-specific immunity-related roles.

In pig, very limited information is available on the MHC class Ib genes *SLA-6*, *-7* and *-8*. The three genes have been fully sequenced from the homozygous Hp1a.0 haplotype [9,10]. Nine allelic variants have been reported for *SLA-6* and only two for *SLA-7* or *SLA-8* [11]. It has been shown that *SLA-Ib* genes are expressed in a less restricted manner than the *HLA-Ib* genes [12, 13] despite a predominant transcription in the lymphoid organs, the lung and the digestive tract [13]. In addition, conversely to the *SLA-Ia* genes, transfection experiments

have revealed that the promoters of *SLA-7* and *SLA-6* do not respond to interferon, suggesting distinct regulatory systems for pig MHC class Ia and Ib genes, as in human [14]. Our aim was to focus on the transcription of the *SLA-7* gene known to have a unique reference transcript [12]. In this report, we show that the *SLA-7* gene expresses a full-length transcript not yet identified as well as at least two additional alternative spliced variants that lead to either exon alteration in the resulting protein or modification of the 3' end of the transcript.

Methods

Animals, tissues and RNA extraction

Tissues from Melanoma-bearing Libechov Minipigs (MeLiM) [15] and French Large White pigs were used. The tissues from MeLiM pigs have been sampled on 13 months old animals. At the time of tissue sampling, all MeLiM animals had regressed, meaning that they were not bearing melanomas anymore [15]. Tissues included brain, thymus, tonsil, spleen and liver. Total RNA was extracted using QIAGEN RNeasy Mini Kits (Qiagen, France). All RNA samples were purified by on-column digestion of DNA with DNase I as recommended by the manufacturer (Qiagen, France).

Primer design

Three primers were designed from the *SLA-7* reference cDNA [12] and genomic [9, 10] sequences, using the Primer3 online program [16]. The primer combinations were suitable to amplify the full coding sequence from exon 1 to the three prime untranslated region (3UTR) or a partial coding sequence from exon 4 to the 3UTR (Table 1 and Figure 1). Primers were

also designed to amplify cDNAs of the *RPL32* gene that has been used as control gene for expression levels (Table 1 and Figure 1).

RT-PCR and sequencing

Two micrograms of DNaseI-treated total RNA were reverse-transcribed (Superscript II enzyme, Invitrogen, USA) with Oligo (dT) primers in a final volume of 20 μ L to which 30 μ L of water were further added to prepare the stock solution of RT samples. PCRs were carried out in a final volume of 15 μ L using 100 nM of each primer, 1 μ L of the 1:10 RT sample and the GoTaqTM DNA polymerase (Promega, USA). Thermocycling conditions were as follows: 94°C for 3 min, followed by 35 amplification cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 90 sec, followed by a final extension at 72°C for 5 min. The PCR products were purified using the JETQUICK Gel Extraction Spin Kit (Genomed, Germany) for further cloning into pCR2.1 vector (TA Cloning Kit, Invitrogen, USA) and sequencing (Eurofins MWG Operon, France).

Sequence analysis

Sequence similarities were searched with the BLAST tools [17]. Multiple alignments were carried out with CLUSTALW [18]. cDNA sequences were translated to protein by online DNA to Protein translation tool (<http://bio.lundberg.gu.se/edu/translat.html>).

Results and discussion

SLA-7 full coding sequences

Full length *SLA-7* transcripts were characterized by RT-PCR from the thymus of MeLiM pigs using the primers *SLA-7-e1-F* and *SLA-7-3UTR-R* (Table 1 and Figure 1). A 1465 nucleotides long transcript was obtained and further named *SLA-7-1465* (Accession number: GU322918). Annotation was carried out by aligning the cDNA sequence to the genomic reference sequence (GenBank accession number AJ251914) and eight exons were detected in this new transcript, in contrast to the reference full-length transcript (Accession number NM_213768) that harbours only seven exons [12] (Figure 1) and is referred to as *SLA-7-001* (OTTSUST00000000782) in the Vertebrate Genome Annotation database [19]. The two RNAs encode proteins that differ in the cytoplasmic tail (Figure 2). The *SLA-7-001* encoded protein contains a cytoplasmic tail that is defined by exons 6 and 7 and is 68 aminoacids long. The *SLA-7-1465* encoded protein is characterized by a cytoplasmic tail that is defined by exons 6 to 8 and is 55 aminoacids long. It has been demonstrated that the cytoplasmic tail of MHC class I molecules contributes to their expression on the cell surface [20] and that mutations of cysteine residues in the cytoplasmic tail of MHC class Ia molecules modify extracellular recognition by Leukocyte Ig-Like receptor 1 [21]. Moreover, it has been reported that HLA-F molecules are entirely dependent on the cytoplasmic tail for export from the endoplasmic reticulum to the Golgi apparatus [22]. Altogether, these reports strongly support a major role for the cytoplasmic tail of MHC class I molecules in transport and function. Further experiments are required to study whether the *SLA-7* molecules encoded by *SLA-7-001* or *SLA-7-1465* transcripts have distinct properties due to their different cytoplasmic tails.

An SLA-7 spliced variant encoding a protein with a shorter alpha 3 domain

A 1366 nucleotides long transcript was retrieved from brain RNA and further referred to as SLA-7-1366 (accession number: HQ224544). Surprisingly, annotation of the cDNA revealed the presence of nine exons due to a splicing site within exon 4 (Figures 1 and 2). The two exons matching to exon 4 were named exons 4a and 4b (Figure 1). Alignment of the SLA-7-1366 cDNA to the reference genomic sequence showed that between exons 4a and 4b, the donor and acceptor splice sites were GA and AG, respectively. This finding indicates that a cryptic non canonical splicing code is used to express this *SLA-7* transcript variant. The general rule is the use of GT and AG for donor and acceptor splicing sites, respectively [23], but alternative codes may be functional [24]. It has been shown that the GA-AG splicing site is rarely used and a few cases have been reported among which splicing in the human *parafibromin* gene [25]. Our results suggest that the *SLA-7* gene may be subject to subtle regulation resulting in the use of rarely used non canonical splicing sites. Additional studies are required to analyze whether this regulation is tissue-specific.

The SLA-7-1366 and SLA-7-1465 encoded molecules with different alpha 3 domain lengths (figures 1 and 2) i.e. 59 (39 from exon 4a and 20 from exon 4b) and 92 aminoacids long, respectively. The alpha 3 domain corresponds to the Immunoglobulin-like region and interacts with the cell surface CD8 glycoproteins that are expressed on cytotoxic T lymphocytes and function as a co-receptor with the T cell receptor [26]. Expression of SLA-7 molecules on the cell surface has not been demonstrated. However, the alpha 3 domain encoded by the SLA-7-1366 transcript is shortened by comparison to the full-length molecule, suggesting that such a modification may alter interactions with cell receptors.

A spliced variant in the 3'UTR with no alteration of the encoded protein

By using primers targeting the three prime end of the gene from exon 4 (SLA-7-e4-F) to the 3UTR (SLA-7-3UTR-R) (Table 1 and figure 1), two partial transcripts were recovered that differ in non coding sequence length (Figure 1). The 650 nucleotide long sequence (SLA-7-650) corresponds to the expected full-length sequence. The 464 nucleotide long sequence SLA-7-464 (accession number: GU322919) is the result of an alternative splicing within the 3UTR region, 31 nucleotides downstream to the stop codon. Modification of the 3UTR does not affect the encoded molecules. As indicated in figure 1, the canonical GT-AG rule was used for this splicing. A new category of transcripts has been recently characterized that can be subject to non sense mediated decay (NMD) [27, 28]. Variants targeted by NMD can present alternative splicing in the 3UTR and the distance between the stop codon and the splice site has been shown to be more than 50 nucleotides long [29]. The SLA-7-464 variant cannot fall into this category of transcripts subject to NMD because the distance between the stop codon and the splicing site is only 31 nucleotides. However, it is tempting to hypothesize that SLA-7-464 variants are subject to a post-transcriptional regulation that has to be explored.

The expression patterns of these two 3UTR variants were studied in four different tissues including spleen, thymus, tonsil, and liver from MeLiM and Large White pigs. Surprisingly, the SLA-7-650 band was detected in all tissues of both breeds but the SLA-7-464 band was detected only in MeLiM pigs (Figure 3). We cannot rule out a very weak expression of the short variant in Large White pigs but our results strongly suggest a co-expression of both variants in MeLiM pigs and a predominant expression of the SLA-7-650 variant in the Large White pigs included in our study.

Conclusion and perspectives

We have identified an *SLA-7* full length transcript that had not been characterized before and that differs from the reference sequence by the length of the encoded cytoplasmic tail. In addition, we show that the *SLA-7* gene is subject to alternative splicing transcription that leads to either a transcript encoding a molecule with a shortened alpha 3 domain or a transcript that is spliced in the 3UTR after the stop codon. In conclusion, the non classical MHC class Ib gene *SLA-7* gene presents a complex transcription pattern, the regulation of which needs to be further investigated. The functions of the putative encoded molecules also need to be studied.

List of abbreviations

MHC : Major Histocompatibility Complex

HLA : Human Leukocyte Antigen

SLA : Swine Leukocyte Antigen

MeLiM : Melanoblastoma-bearing Libechov Minipigs

3UTR : 3 prime untranslated region

Class Ib : non classical class I

Class Ia : classical class I

CDS: coding sequence

NMD : Non sense Mediated Decay

PCR : Polymerase Chain Reaction

RT : reverse transcription

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RH participated in working out the experimental design, performed all experiments and sequence analysis, and drafted the manuscript. GL carried out RT-PCR experiments. EB and SVN contributed to tissue sampling. CRG coordinated the whole study, contributed to the experimental design, the analysis and interpretation of the results and corrected the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

The authors would like to thank Dr. Patrick Chardon (GABI-GIS, INRA, France) for useful discussion and H  l  ne Hayes for English correction of the manuscript. Rui Hu's PhD was supported by the Animal Genetics Department of INRA and Dadi Agriculture Comprehensive Exploitation Co. Ltd, Liaoning, China.

References

1. Townsend A, Bodmer H: **Antigen recognition by class I-restricted T lymphocytes.** *Annu Rev Immunol* 1989, **7**:601-624.
2. Trowsdale J, Barten R, Haude A, Stewart CA, Beck S, Wilson MJ: **The genomic context of natural killer receptor extended gene families.** *Immunol Rev* 2001, **181**:20-38.
3. Moscoso J, Serrano-Vela JJ, Pacheco R, Arnaiz-Villena A: **HLA-G, -E and -F: allelism, function and evolution.** *Transpl Immunol* 2006, **17**(1):61-64.
4. Carosella ED, Moreau P, Le Maoult J, Le Discorde M, Dausset J, Rouas-Freiss N: **HLA-G molecules: from maternal-fetal tolerance to tissue acceptance.** *Adv Immunol* 2003, **81**:199-252.
5. Joly E, Rouillon V: **The orthology of HLA-E and H2-Qa1 is hidden by their concerted evolution with other MHC class I molecules.** 2006, *Biol Direct* **1**: 2.
6. Comiskey M., Goldstein C.Y., De Fazio S.R., Mammolenti M., Newmark J.A. & Warner C.M. (2003): **Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product.** *Hum Immunol* 64, 999-1004.
7. Lau P, Amadou C, Brun H, Rouillon V, McLaren F, Le Rolle AF, Graham M, Butcher GW, Joly E: **Characterisation of RT1-E2, a multigenic family of highly conserved rat non-classical MHC class I molecules initially identified in cells from immunoprivileged sites.** *BMC Immunology* 2003, **4**:7.

8. Birch J, Codner G, Guzman E, Ellis SA: **Genomic location and characterisation of nonclassical MHC class I genes in cattle.** *Immunogenetics* 2008, **60**(5):267-273.
9. Chardon P, Rogel-Gaillard C, Cattolico L, Duprat S, Vaiman M, Renard C: **Sequence of the swine major histocompatibility complex region containing all non-classical class I genes.** *Tissue Antigens* 2001, **57**(1):55-65.
10. Renard C, Hart E, Sehra H, Beasley H, Coggill P, Howe K, Harrow J, Gilbert J, Sims S, Rogers J, Ando A, Shigenari A, Shiina T, Inoko H, Chardon P, Beck S: **The genomic sequence and analysis of the swine major histocompatibility complex.** *Genomics* 2006, **88**: 96-110.
11. Ho CS, Lunney JK, Ando A, Rogel-Gaillard C, Lee JH, Schook LB, Smith DM: **Nomenclature for factors of the SLA system, update 2008.** *Tissue Antigens* 2009, **73**: 307-15.
12. Crew MD, Phanavanh B, Garcia-Borges CN: **Sequence and mRNA expression of nonclassical SLA class I genes SLA-7 and SLA-8.** *Immunogenetics* 2004, **56**(2):111-114.
13. Kusza S, Flori L, Gao Y, Teillaud A, Hu R, Lemonnier G, Bősze Z, Bourneuf E, Vincent-Naulleau S, Rogel-Gaillard C: **Transcription specificity of the class Ib genes SLA-6, SLA-7 and SLA-8 of the swine major histocompatibility complex and comparison with class Ia genes.** *Anim Genet* 2010, in press
14. Tennant, L.M. *et al*: **Regulation of porcine classical and nonclassical MHC class I expression.** *Immunogenetic* 2007, **59**(5): 377-89.

15. Vincent-Naulleau S, Le Chalony C, Leplat JJ, Bouet S, Bailly C, Spatz A, Vielh P, Avril MF, Tricaud Y, Gruand J *et al*: **Clinical and histopathological characterization of cutaneous melanomas in the melanoblastoma-bearing Libechov minipig model.** *Pigment Cell Res* 2004, **17**(1):24-35.
16. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers.** *Methods Mol Biol* 2000, **132**:365-386.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**(3):403-410.
18. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**(22):4673-4680.
19. **VEGA** [<http://vega.sanger.ac.uk/index.html>]
20. Davis DM, Reyburn HT, Pazmany L, Chiu I, Mandelboim O, Strominger JL: **Impaired spontaneous endocytosis of HLA-G.** *Eur J Immunol* 1997, **27**(10):2714-2719.
21. Gruda R, Achdout H, Stern-Ginossar N, Gazit R, Betser-Cohen G, Manaster I, Katz G, Gonen-Gross T, Tirosh B, Mandelboim O: **Intracellular cysteine residues in the tail of MHC class I proteins are crucial for extracellular recognition by leukocyte Ig-like receptor 1.** *J Immunol* 2007, **179**(6):3655-3661.

22. Boyle LH, Gillingham AK, Munro S, Trowsdale J: **Selective export of HLA-F by its cytoplasmic tail.** *J Immunol* 2006, **176**(11):6464-6472.
23. Mount SM: A catalogue of splice junction sequences. *Nucleic Acids Res* 1982, **10**(2):459-72.
24. Burset M, Seledtsov IA, Solovyev VV: **Analysis of canonical and non-canonical splice sites in mammalian genomes.** *Nucleic Acids Res* 2000, **28**(21):4364-4375.
25. Bradley KJ, Cavaco BM, Bowl MR, Harding B, Young A, Thakker RV: **Utilisation of a cryptic non-canonical donor splice site of the gene encoding PARAFIBROMIN is associated with familial isolated primary hyperparathyroidism.** *J Med Genet* 2005, **42**(8):e51.
26. Sun J, Kavathas PB: **Comparison of the roles of CD8 alpha alpha and CD8 alpha beta in interaction with MHC class I.** *J Immunol* 1997, **159**(12):6077-6082.
27. Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ: **Deciphering the splicing code.** *Nature* 2010, **465**(7294):53-59.
28. Gardner LB: **Nonsense-mediated RNA decay regulation by cellular stress: implications for tumorigenesis.** *Mol Cancer Res* 2010, **8**(3):295-308.
29. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC: **Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise.** *Nat Genet* 2004, **36**(10):1073-1078.

Table 1. RT-PCR primers

Genes	Primer name	Sequence of primers (5'-3')	Position of primers	Accession numbers
SLA-7	SLA-7-e1-F	ATGGGGCCCCGAGCCCTCCTCCT	Exon1	AJ251914 AY463541
	SLA-7-e4-F	TGGAGAGGAGCAGAGCTACA	Exon4	
	SLA-7-3UTR-R	AGAGCCACTGCTGATCCAGT	3'UTR	
RPL32	RPL32-F	TGCTCTCAGACCCCTTGTGAAG	Exon1	NM_001001636
	RPL32-R	TTTCCGCCAGTTCCGCTTA	Exon2	

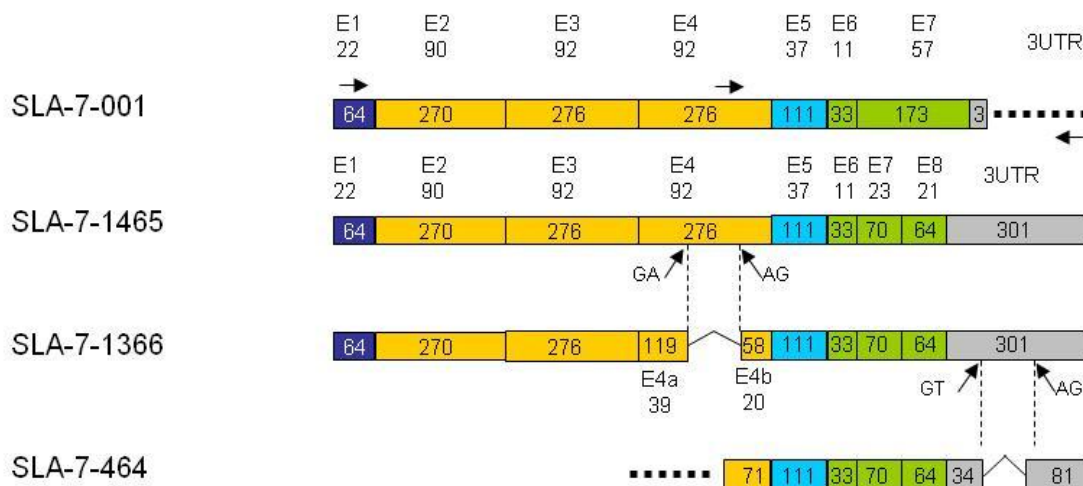


Figure 1. Schematic representation of the *SLA-7* transcripts.

Exons are numbered E1 to E8 and the three prime non coding region is indicated as 3UTR. Sizes of exons and 3UTR are given in nucleotides within the boxes. The number of aminoacids for each exon is indicated above the exon number. Exons represented by a dark blue box (E1) correspond to the leader sequences. Exons represented by orange boxes (E2 to E4) stand for the alpha1, 2 and 3 domains of the molecules. Exons represented in bright blue (E5) correspond to the transmembrane domain. Exons represented by green boxes encode the cytoplasmic tail of the molecule. The 3UTR is represented by a grey box. Positions of the primers used for RT-PCRs are indicated by arrows on top of the figure on E1 (SLA-7-e1-F), E4 (SLA-7-e4-F) and 3UTR (SLA-7-3UTR-R). The donor and acceptor splice sequences are positioned by arrows on E4 and 3UTR boxes.

```

SLA-7-1366      MGPOALLLLLSGTLVLTQFWARPHSLRYFYTAVSRPSRRDPRFSVVGYYDDTQFVRFDS 60
SLA-7-1465      MGPOALLLLLSGTLVLTQFWARPHSLRYFYTAVSRPSRRDPRFSVVGYYDDTQFVRFDSN 60
SLA_7_001       MGPRALLLLLSGTLVLTQFWARPHSLRYFYTAVSRPSRRDPRFSVVGYYDDTQFVRFDSN 60
***-*****_

SLA-7-1366      APNPREEPRTFWVELEGPEYCDNRNTRIYKDTSONFQVSLNVLRGYYNQSEAESHYQWLC 120
SLA-7-1465      APNPREEPRTFWVELEGPEYCDNRNTRIYKDTSONFQVSLNVLRGYYNQSEAESHYQWLC 120
SLA_7_001       APNPREEPRTFWVELEGPEYCDNRNTRIYKDTSONFQVSLNVLRGYYNQSEAESHYQWLC 120
*****_*****

SLA-7-1366      GCYVARDGRLLRGYSQFAYDGADYIALNEDLRSWTAVDMAAQITRRKWEFEETVAEQSRAY 180
SLA-7-1465      GCYVARDGRLLRGYSQFAYDGADYIVLNEDLRSWTAVDMAAQITRRKWEFEETVAEQSRAY 180
SLA_7_001       GCYVARDGRLLRGYSQFAYDGADYIVLNEDLRSWTAVDMAAQITRRKWEFEETVAEQSRAY 180
*****_*****

SLA-7-1366      LEVACVQSLHRYLANGKETLQRSPPKTHATRHPSSDNKVTLRCWALGFYPKEISLTWQQ 240
SLA-7-1465      LEVACVQSLHRYLVNGKETLQRSPPKTHVTRHPSSDNKVTLRCWALGFYPKEISLTWQQ 240
SLA_7_001       LEVACVQSLHRYLANGKETLQRSPPKTHVTRHPSSDNKVTLRCWALGFYPKEISLTWQQ 240
*****_*****

SLA-7-1366      EGQ-----SYTCHVQOEGLOESLTLRWDPQP 267
SLA-7-1465      EGQDQSQDVEVVETRPSGDGTFRWAALVVPPEEQSYTCHVQOEGLOESLTLRWDPQP 300
SLA_7_001       EGQDQSQDVEVVETRPSGDGTFRWAALVVPPEEQSYTCHVQOEGLOESLTLRWDPQP 300
***                      *****

SLA-7-1366      PVPIVGIIVGLVIVLVAGAVVTGVVIWRKKCSGGKVRKYQQAESSRSENSGVCIMPFKA 327
SLA-7-1465      PVPIVGIIVGLVIVLVAGAVVTGVVIWRKKCSGGKVRKYQQAESSRSENSGVCIMPFKA 360
SLA_7_001       PVPIVGIIVGLVIVLVAGAVVTGVVIWRKKCSGGKVRKYQQAESSRSENSGVCIMPFKA 360
*****

SLA-7 1366      ETLEALSGKQLPYTRRTD-----QVTNISIMAL----- 355
SLA-7-1465      ETLEALSGKQLPYTRRTD-----QVTNISIMAL----- 388
SLA_7_001       ETLEALRWDSGWGRGDTTGLWGFKWNLSMAGAVRRSALM 401
*****  ..      *      :  .*: :  .

```

Figure 2. Multi-alignment of peptides encoded by the transcripts SLA-7-1465, SLA-7-1366 and SLA-7-001. The successive eight (SLA-7-1366 and SLA-7-1465) or seven (SLA-7-001) exons are alternatively indicated by black and blue font. Aminoacids at the junction between two exons are in grey boxes. Aminoacid similarities between two or three sequences are indicated below the sequence alignments by dots or stars, respectively.

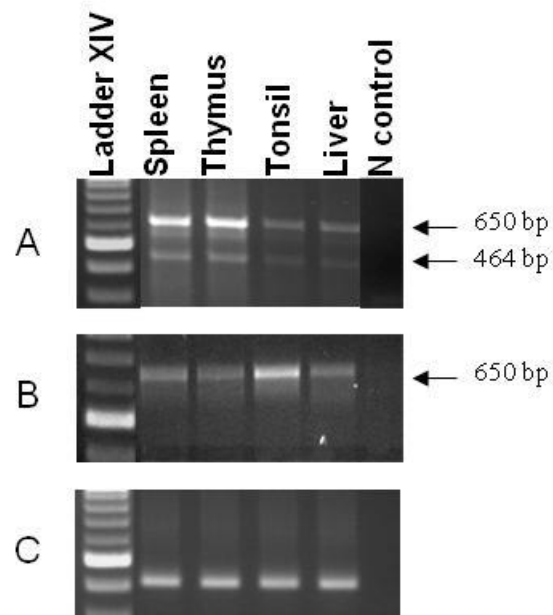


Figure 3. Tissue expression patterns in MeLiM and Large White pigs.

Detection of the *SLA-7* transcript variants in adult tissues from MeLiM (A) and Large White (B) pigs by RT-PCR using the primers *SLA-7-e4-F* and *SLA-7-3UTR-R*. The *RPL32* gene was used as a control for expression levels as shown for four tissues of MeLiM pig (C).

Publication # 2

Animal Genetics, in press

Transcription specificity of the non classical class I genes SLA-6, SLA-7 and SLA-8 of the swine major histocompatibility complex and comparison with classical class I genes

Szilvia Kusza, Laurence Flori, Yu Gao, Angélique Teillaud, **Rui Hu**, Gaëtan Lemonnier, Zsuzsanna Bősze, Emmanuelle Bourneuf, Silvia Vincent-Naulleau, Claire Rogel-Gaillard

Transcription specificity of the class Ib genes *SLA-6*, *SLA-7* and *SLA-8* of the swine major histocompatibility complex and comparison with class Ia genes

S. Kusza*, L. Flori^{†,‡,§}, Y. Gao^{†,‡,§}, A. Teillaud^{†,‡,§}, R. Hu^{†,‡,§}, G. Lemonnier^{†,‡,§}, Z. Bösze[¶], E. Bourneuf^{†,‡,§}, S. Vincent-Naulleau^{†,‡,§} and C. Rogel-Gaillard^{†,‡,§}

*Institute of Animal Science, University Debrecen, Centre of Agricultural Sciences and Engineering, 4032 Debrecen, Böszörményi Str.138, Hungary. [†]INRA, UMR 1313 de Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France. [‡]CEA, DSV, IRCM, SREIT, Laboratoire de Radiobiologie et Etude du Génome, 78350 Jouy-en-Josas, France. [§]AgroParisTech, UMR 1313 de Génétique Animale et Biologie Intégrative, 78350 Jouy-en-Josas, France. [¶]Agricultural Biotechnology Center, 2100 Gödöllő, Szent Györgyi A. Str. 4, Hungary

Summary

Our aim was to analyse the transcription levels of the three non-classical class Ib genes *SLA-6*, *SLA-7* and *SLA-8*, of the swine major histocompatibility complex, in various tissues and conditions and to compare them to the transcription levels of classical class Ia genes. Twenty-five adult tissues from two pig breeds, pig renal PK15 cells infected with the Pseudorabies virus and peripheral blood mononuclear cells (PBMCs) stimulated by lipopolysaccharide or a mixture of phorbol myristate acetate and ionomycin were included in our study. Relative transcription was quantified by quantitative real-time PCR. On average, in adult tissues and PBMCs and compared to *SLA-6*, the transcription level of *SLA-Ia* genes was 100–1000 times higher, the level of *SLA-8* was 10–20 times higher, and that of *SLA-7* was five times higher. Thus, *SLA-8* is the most transcribed *SLA-Ib* gene, followed by the *SLA-7* and *SLA-6* genes. The highest transcription levels of *SLA-Ib* transcripts were found in the lymphoid organs, followed by the lung and the digestive tract. The tissue variability of expression levels was widest for the *SLA-6* gene, with a 1:32 ratio between the lowest and highest levels in contrast to a 1:12 ratio for the *SLA-7* and *SLA-8* genes and a 1:16 ratio for the *SLA-Ia* genes. During PK-15 infection and PBMC stimulation, *SLA-Ia* and *SLA-8* genes were down-regulated, whereas *SLA-6* and *SLA-7* were up-regulated, down-regulated or not significantly modified. Our overall results confirm the tissue-wide transcription of the three *SLA-Ib* genes and suggest that they have complementary roles.

Keywords class Ia, class Ib, major histocompatibility complex, pig, *SLA-6*, *SLA-7*, *SLA-8*, transcription.

Introduction

The major histocompatibility complex (MHC) locus is a large and highly gene dense region that contains many genes involved in adaptive and innate immune responses and inflammation, as well as numerous other genes, the functions of which have not all been identified (Horton *et al.* 2004; Kelley *et al.* 2005). In particular, the MHC locus contains a large multigene family that encodes the membrane-bound glycoproteins known as the histocompatibility

or MHC molecules, which play a major role in peptide recognition and presentation as well as in graft rejection (Horton *et al.* 2004; Kelley *et al.* 2005). Among these molecules, the classical MHC class I molecules (MHC class Ia) present peptides derived either from self-proteins or from proteins of intracellular pathogens to CD8⁺ cytotoxic T cells (Rock *et al.* 2010). They bind to inhibitory receptors on NK cells that include killer cell Ig-like receptors (KIR) in man, C-type lectin-like Ly49 molecules in mouse and CD94/NKG2A heterodimers in man and mouse (Lanier 2001; Anfossi *et al.* 2006). MHC class Ia genes are extremely polymorphic, and their expression has been detected in most cell types (Shawar *et al.* 1994; Rodgers & Cook 2005). Conversely, non-classical MHC class I (MHC class Ib) genes such as the human genes *HLA-E*, *HLA-F* and *HLA-G* are

Address for correspondence

✉ E-mail: claire.rogel-gaillard@jouy.inra.fr

Accepted for publication 10 October 2010

Animal Genetics © 2010 Stichting International Foundation for Animal Genetics

No claim to original US government works

	A G E	2 1 7 0	B	Dispatch: 29.12.10	Journal: AGE	CE: Anusha
Journal Name	Manuscript No.			Author Received:	No. of pages: 11	PE: Prasanna

oligomorphic and show tissue-specific expression (Shawar *et al.* 1994; Rodgers & Cook 2005). HLA-E, HLA-F and HLA-G molecules can associate with beta-2 microglobulin (B2M), HLA-G and HLA-E molecules bind T-cell receptors (Rodgers & Cook 2005), HLA-G molecules bind inhibitory NK-cell receptors (LIRs), and HLA-E molecules interact with inhibitory (CD94/NKG2A heterodimers) and activating (CD94/NKG2C heterodimers) NK-cell receptors (Rodgers & Cook 2005). These interactions with NK-cell receptors are involved in immunomodulation, allergy, autoimmunity and embryonic development (LeMaout *et al.* 2005; Rodgers & Cook 2005). HLA-E, HLA-F and HLA-G molecules probably play distinct and complementary functions. HLA-G inhibits the cytolytic functions of NK cells and cytotoxic T lymphocytes and plays a key role in foeto-maternal tolerance during pregnancy (Hunt *et al.* 2007) and establishment of immune tolerance in tumorigenesis (Gomes *et al.* 2007; Carosella *et al.* 2008). HLA-E presents peptides derived from the leader sequence of MHC class I molecules as well as other peptides derived from pathogens like the cytomegalovirus, thus providing evidence for a role in bridging innate and adaptive immune responses (Sullivan *et al.* 2008). In contrast to HLA-E and HLA-G, HLA-F seems to act independently on peptide presentation and the encoded molecules are predominantly expressed inside the cells (Wainwright *et al.* 2000; Boyle *et al.* 2006; Apps *et al.* 2008). Peptide presentation by HLA-F molecules has not been reported but is not ruled out (Rodgers & Cook 2005). Orthology between *HLA-E* and the mouse gene *H2-Qa1* has been confirmed (Joly & Rouillon 2006), but to date, no orthologs have been identified for *HLA-F* and *HLA-G* despite functional homologies between *HLA-G* and *H2-Qa2* (Comiskey *et al.* 2003; Gomes *et al.* 2007). Four MHC non-classical class I genes have been identified in cattle, but neither sequence orthology nor functional homologies with human or mouse genes have been found so far (Birch *et al.* 2008).

In pig, the MHC locus is a 2.4-megabase region that has been extensively sequenced and annotated (Renard *et al.* 2006). Three classical class Ia genes *SLA-1*, *SLA-2* and *SLA-3* have been identified as functional, and classification of alleles is available (Smith *et al.* 2005; Ho *et al.* 2009). It is commonly assumed that the highly polymorphic SLA class Ia genes are functional orthologs of the human MHC class Ia genes *HLA-A*, *HLA-B* and *HLA-C*. The SLA class Ia molecules have been shown to be directly recognized by subpopulations of both human CD8⁺ cytolytic T cells (Shishido *et al.* 1997; Xu *et al.* 1999) and NK cells (Donnelly *et al.* 1997; Itescu *et al.* 1998), leading to the lysis of the target cells. The three genes *SLA-6*, *SLA-7* and *SLA-8* are classified as non-classical class I genes because of their limited polymorphism and slight variations in the three prime end specific of the cytoplasmic tail compared to the *SLA-1*, *SLA-2* and *SLA-3* genes (Chardon *et al.* 2001). In contrast to SLA-Ia genes, no orthology or functional homology has ever been established with HLA class Ib genes, and neither

gene mapping nor sequence phylogeny is helpful in this case (Chardon *et al.* 2001; Crew *et al.* 2004; Renard *et al.* 2006; Lunney *et al.* 2009). Addressing the role of SLA class Ib genes in pig is a major issue given the role of HLA class Ib genes in immunotolerance (Moscoso *et al.* 2006). In addition, the SLA-Ib genes could be major candidate genes for species-specific functions. Although SLA-Ib genes have been characterized at the sequence level (Chardon *et al.* 2001), little is known about their expression and specificity, and no antibody is suitable for specific detection of *SLA-6*, *SLA-7* or *SLA-8* proteins. It has been reported that SLA-Ib genes are transcribed in a less restricted manner than HLA class Ib genes, but the number and variety of tissues included in this unique study were limited (Crew *et al.* 2004). As a first step towards functional studies on SLA-Ib genes, our aim was to analyse the transcription levels of the three genes *SLA-6*, *SLA-7* and *SLA-8* in a wide range of adult pig tissues and under different immune stimulation conditions and to compare the transcription levels to those of classical genes under the same conditions.

Material and methods

Animals, tissues and cells

Tissues were collected from three healthy Melanoma bearing Lieberman Minipigs (MeLiM) (Horak *et al.* 1999; Vincent-Naulleau *et al.* 2004) and two French Large White pigs. Tissues were frozen in liquid nitrogen immediately after sampling for further storage at -80 °C. Blood was sampled with sodium heparin-coated tubes from seven 60-day-old castrated Large White males, and peripheral blood mononuclear cells (PBMCs) were purified by Ficoll gradient and either mock-stimulated (C) or stimulated for 24 h with lipopolysaccharide (LPS) at 1 µg/ml (L) or a mixture of phorbol myristate acetate (PMA) at 10 ng/ml and ionomycin at 1 µg/ml (P) under conditions reported elsewhere (Gao *et al.* 2010). The renal epithelial cells PK15 were cultured in synthetic medium and either mock-infected (MI_T0) or infected by the virulent strain NIA3 of the Pseudorabies virus (PrV) under previously reported conditions (Flori *et al.* 2008a,b). Three biological replicates were produced, and PK15 cells were collected at 0, 2, 4, 8 and 12 (I_T0, I_T2, I_T4, I_T8 and I_T12) hours post-infection.

RNA extraction and quality control

Total RNA was extracted from frozen adult tissues or freshly collected cells. Commercial kits (Qiagen) were chosen according to tissue specificity: (1) the RNeasy Midi kit was used for duodenum, ileum, jejunum, kidney, liver, lung, Peyer's patches, spleen, thymus, tonsil, cornea, epididymis, testis, ovary, uterus, adrenal gland, inguinal lymph node and PBMCs; (2) the RNeasy Fibrous Tissue kit was chosen for diaphragma, longissimus dorsi, heart (ventricle), nasal

mucosa and aorta; (3) the RNeasy Lipid Tissue kit was used for back fat, skin and brain. RNA extractions were carried out following the manufacturer's instructions. Finally, for PK15 cells, total RNA was extracted with the Trizol-chloroform method (Invitrogen, France). For all samples, residual genomic DNA was removed by on-column digestion of DNA with DNase I (DNase set and cleanup, Qiagen). RNA concentration was determined by Nanodrop quantification (Thermo Fisher Scientific Inc.). RNA quality was checked on an Agilent 2100 Bioanalyzer (Agilent Technologies). RNAs with a RIN score between 8 and 10 were used. The concentration of RNA stock solutions was adjusted at 1 µg/µl in water for storage at -80 °C.

PCR primers design and sequence alignment

PCR primers were selected using the Primer Express software (Applied Biosystem). Sequence multi-alignments were carried out with the clustal W program (Larkin *et al.* 2007).

SLA-Ia and SLA-Ib genes share high sequence similarities. To design primers in sequence segments harbouring gene-specific nucleotide variations, sequences were multi-aligned before primer selection. For SLA-1, SLA-2 and SLA-3 genes, available alleles were multi-aligned to identify non-polymorphic sequences and further multi-aligned to SLA-Ib cDNAs to identify SLA-Ia-specific monomorphic sequences. In Fig. 1, multi-alignments of cDNAs were produced with the full-length cDNAs published by Crew and collaborators (Crew *et al.* 2004) for SLA-6, SLA-7 and SLA-8 and with available full-length cDNAs for SLA-1 (AY247766), SLA-2 (DQ104339) and SLA-3 (AY135602). Primers overlapping exons 2 and 3 were chosen for each SLA-6, SLA-7 and SLA-8 gene. A unique primer pair spanning exons 5 and 6 was designed to simultaneously amplify the classical genes

SLA-1, SLA-2 and SLA-3 (Table 1 and Fig. 1). The primers targeting the three functional SLA-Ia gene transcripts were further referred to as SLA-Ia primers. For SLA-6, SLA-7 and SLA-8 genes, primers were derived from the reference sequence of the Hpl.a.0 haplotype (Renard *et al.* 2006).

Reverse transcription (RT), RT-PCR and qRT-PCR

Two and a half micrograms of DNaseI-treated total RNA were reverse-transcribed using Superscript II enzyme with Oligo dT₁₂₋₁₈ (Invitrogen, Carlsbad, CA) and random primers (Promega) as recommended by the enzyme manufacturer (Invitrogen, Carlsbad, CA). RT reactions were carried out in a final volume of 30 µl with RNase OUT (20 units). RT reactions were adjusted to 50 µl with water and then considered as non-diluted RT products. RT-PCRs were carried out using either non-diluted or various dilutions of the RT products as DNA templates. For all primer pairs, PCR thermocycling conditions were as follows: initial denaturation at 94 °C for 30 s followed by 35 cycles of the three steps 94 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s and a final storage at 10 °C. For qRT-PCR, the cDNAs were quantified using a 2100 Bioanalyzer (Agilent Technologies) and diluted to a working concentration of 4 ng/µl. Duplicate (adult tissues, PBMCs) or triplicate (PK15 cells) reactions were performed in a final volume of 20 µl with 20 ng cDNA, 300 nm primers and SYBR Green PCR Master Mix (Applied Biosystem), using an ABI PRISM 7900 HT sequence detection system (Applied Biosystem). Five housekeeping genes (B2M, RPL32, GAPDH, HPRT1, ACTB) were tested with the Genorm software (Vandesompele, 2002 #327), which indicated that the genes B2M and RPL32 were equally good reference genes. The gene B2M was chosen as an internal reference gene for Large White

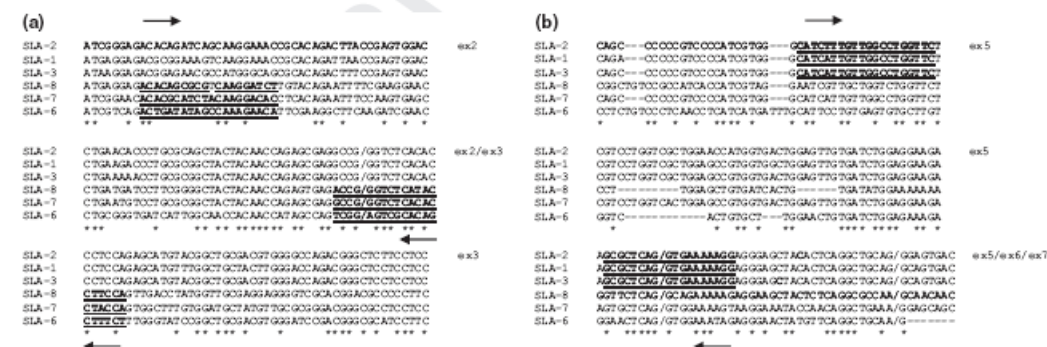


Figure 1 Sequence multi-alignment of SLA-Ia and SLA-Ib cDNAs and primer position. (a). Partial multi-alignment of exons 2 and 3 and positions of primers targeting SLA-6, SLA-7 and SLA-8 cDNAs. The SLA-7 reverse primer is common to SLA-Ia cDNAs, but the direct primer shows mismatches suitable for specific amplification of SLA-7. The SLA-8 direct primer was derived from the sequence of the Hpl.a.0 haplotype and shows one mismatch with the aligned sequence (T for the aligned sequence and C for SLA-8 of the Hpl.a.0 haplotype). (b). Partial multi-alignment of exons 5-7 and positions of primers targeting the three SLA-Ia cDNAs. The direct primer is common to SLA-7 but the reverse primer shows mismatches suitable for specific amplification. cDNA sequence accession numbers are AY247766 for SLA-1, DQ104339 for SLA-2, AY135602 for SLA-3, AY463540 for SLA-6, AY463541 for SLA-7 and AY463542 for SLA-8. All primer information is summarized in Table 1. Arrows indicate primer orientation.

Animal Genetics © 2010 Stichting International Foundation for Animal Genetics

No claim to original US government works, doi: 10.1111/j.1365-2052.2010.02170.x

Table 1 Gene specific primers.

Genes	Primer name	Sequence of primers (5'-3')	Position of primers	Accession numbers	Amplicon length (bp)
SLA-6	SLA6_e2_F	ACTGATATAGCCAAAGAACA	Exon 2	AJ251914 ¹	98
	SLA6_e23_R	AGAAAGCTGTGCGACTCCGA	Exons 2/3 overlap		
SLA-7	SLA7_e2_F	ACACGCATCTACAAGGACAC	Exon 2	AJ251914 ¹	98
	SLA7_e23_R	TGGTAGGTGTGAGACCCGGC	Exons 2/3 overlap		
SLA-8	SLA8_e2_F	ACACAGCGCGCCAAGGATCT	Exon 2	AJ251914 ¹	98
	SLA8_e23_R	TGGAAGGTATGAGACCCGGT	Exon 2/3		
SLA-Ia ²	SLAIa_e5_F	CATCATTGTTGGCCTGGTTC	Exon 5	AY247766, DQ104339, AY135602	89
	SLAIa_e56_R	CCTTTTTCACCTGAGCGC	Exons 5/6 overlap		
B2M	B2M_F	TGGTCTTCTACCTTCTGGTCC	Exon 2	NM_213978	134
	B2M_R	TGTGATGCCGGTAGTGGTCTC	Exon 3		
RPL32	RPL32_F	TGCTCTCAGACCCCTTGGAAG	Exon 1	NM_001001636	106
	RPL32_R	TTTCCGCCAGTTCCGCTTA	Exon 2		

¹The primers were derived from the genomic sequence corresponding to the haplotype Hp1a.0.

²The primers were designed to co-amplify SLA-1, SLA-2 and SLA-3 cDNAs.

adult tissues, and stimulated or mock-stimulated PBMCs and RPL32 was chosen as the internal reference gene for adult MeLiM tissues and mock-infected or infected PK15 cells. The $2^{-\Delta\Delta Ct}$ method was used to calculate the fold change in gene expression. For the relative quantification of each gene according to tissues, the brain was used as calibrator. For the relative quantification of genes per tissue, SLA-6 transcription levels were used as calibrators. Significance of the difference in average transcription levels between genes and conditions was tested using a two-sided paired *t*-test (R software, <http://www.r-project.org/>). Spearman's rank correlation coefficient between SLA-Ia, SLA-6, SLA-7 and SLA-8 transcription levels and their significance were calculated with the *cor.test* function (stats package, R software).

Results

Primer specificity

Primers were designed to have a primer pair specific for each SLA-6, SLA-7 and SLA-8 gene and a unique primer pair targeting the three classical genes SLA-1, SLA-2 and SLA-3. SLA-7 cDNA and SLA-Ia sequence share a higher similarity level than SLA-6 and SLA-8 cDNAs. The SLA-7 reverse primer shows only one mismatch with exon 3 of SLA-Ia cDNAs (Fig. 1A), but the direct primer shows more mismatches suitable for specific amplification of SLA-7. Similarly, the SLA-Ia direct primer fully matches with SLA-7 cDNA at the fifth position downstream from the five prime end (see Fig. 1B), but the SLA-Ia reverse primer shows one mismatch with the SLA-7 sequence at the three prime end, which is suitable for specific amplification of SLA-Ia cDNAs. All PCR products were checked by sequencing (data not shown). The SLA primers are expected to be non-haplotype-

specific and to efficiently amplify cDNAs from Large White and MeLiM pig tissues (Fig. 2) as well as cDNAs from the renal PK15 cell line. All SLA as well as B2M and RPL32 reference gene primers were shown to be efficient to carry out RT-PCRs (Fig. 2) as well as qRT-PCR experiments.

Relative transcription levels of SLA-Ia, SLA-6, SLA-7 and SLA-8 genes

qRT-PCRs were carried out for 25 distinct adult tissues from either three MeLiM or two Large White pigs, for PBMCs from seven Large White pigs either mock-stimulated or stimulated by LPS or a mixture of PMA and ionomycin, and for epithelial PK15 cells during the time course of infection by the PrV. Quantification of the relative transcription levels of SLA-Ia and SLA-Ib genes was performed using SLA-6 normalized data as calibrators (Table 2). Significant differences in relative average transcription levels were found between SLA-Ia, SLA-7 and SLA-8 genes in MeLiM and Large White tissues (*P*-value $< 6 \times 10^{-5}$). Transcription levels of SLA-Ia genes were 74 (in kidney) to 1453 (in testis) times higher than those of SLA-6 in adult MeLiM tissues and 31 (in thymus) to 2437 (in nasal mucosa) times higher than those of SLA-6 in adult Large White tissues. Expression levels of SLA-7 were 1.5 (in inguinal lymph node) to 18 (in testis) times higher than those of SLA-6 in MeLiM adult tissues and up to eight times higher in the adrenal gland of Large White pigs when compared to SLA-6. The transcription level of SLA-8 was 4 (in thymus) to 89 (in testis) times higher than that of SLA-6 in MeLiM tissues and 2 (in spleen) to 33 (in brain) times higher than that of SLA-6 in Large White tissues. On average, transcription levels of SLA-Ia were 530–632 times higher than those of SLA-6 in the adult tissues of MeLiM and Large White pigs, respectively. Compared to SLA-6, the average transcription levels of SLA-7 and SLA-8 were, respectively, 5 and 23 times higher in MeLiM

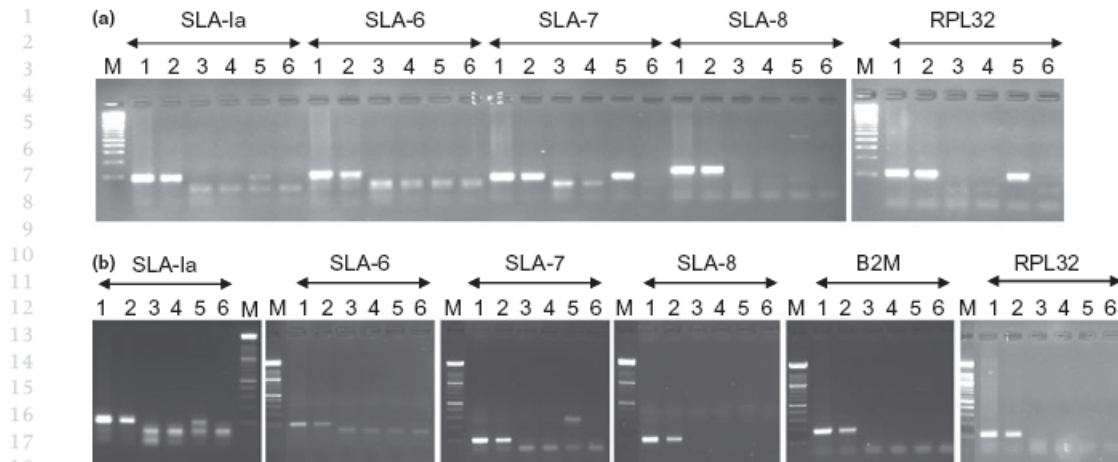


Figure 2 RT-PCR results using SLA-Ia-, SLA-6-, SLA-7-, SLA-8-, B2M- and RPL32-specific primers. Results are presented for the MeLiM pig 484 using kidney cDNA (a) and for a Large White pig using cDNA from dermal fibroblasts (b). M: DNA ladder, 1: non-diluted positive RT products, 2: 1/4 diluted positive RT products, 3: non-diluted negative RT products, 4: 1/4 diluted negative RT products, 5: genomic DNA, 6: water.

tissues and 4 and 12 times higher in Large White tissues. These results clearly show that the SLA-Ia genes are significantly more transcribed than the SLA-Ib genes. In addition, among the three non-classical genes, SLA-8 and SLA-6 are the most and least transcribed, respectively. This order was confirmed in stimulated and non-stimulated PBMCs as well as in PK15 cells infected by PrV or mock-infected (Table 2). On average, in PBMCs, compared to the transcript levels of SLA-6, those of SLA-Ia, SLA-7 and SLA-8 were 928 times higher, equivalent (or slightly lower) and 22 times higher, respectively. On average in PK15 cells, compared to transcript levels of SLA-6, those of SLA-Ia, SLA-7 and SLA-8 were, respectively, 11, 2 and 4 times higher. All results obtained either from adult tissues, PBMCs or PK15 cell lines indicate SLA-Ia to be the most transcribed gene, followed by SLA-8, SLA-7 and finally SLA-6. The order of magnitude is approximately 1 to 10–20 between SLA-6 and SLA-8 in all tested tissues and cells and between 1 to 100–1000 between SLA-6 and SLA-Ia in adult tissues and PBMCs. However, the ratio is closer to 1:10 between SLA-6 and SLA-Ia for the cell line PK15. Finally, in the tissues and conditions included in this study, the ratio between SLA-7 and SLA-6 did not exceed 5. In addition, SLA-Ia, SLA-7 and SLA-8 transcription levels were moderately correlated in Large White tissues ($\rho = 0.6$, P -value $< 10^{-3}$) and highly correlated in MeLiM tissues ($\rho > 0.75$, P -value $< 3 \times 10^{-5}$).

Variation in *in vivo* transcription levels of SLA-Ia, SLA-6, SLA-7 and SLA-8 according to adult tissues

For each gene, the relative transcription levels in adult tissues were calculated for MeLiM animals using the

transcription level found in brain as calibrator (Fig. 3). They ranged between 0.5 and 8 for SLA-Ia, 0.5 and 6 for SLA-7 and SLA-8 and between 0.5 and 16 for SLA-6. Thus, the ratios between the lowest and the highest transcription levels among tissues were 1:32 for SLA-6, 1:16 for SLA-Ia and 1:12 for SLA-7 and SLA-8. Transcription levels of the SLA-Ia and the three non-classical genes were positively correlated, and correlations were high with SLA-7 and SLA-8 ($\rho = 0.78$, P -value = 5×10^{-5}) and to a lesser extent with SLA-6 ($\rho = 0.59$, P -value = 3×10^{-3}). For the three non-classical gene transcription levels, the correlation coefficient was 0.76 between SLA-6 and SLA-8 (P -value = 5×10^{-5}), 0.76 between SLA-6 and SLA-7 (P -value = 5×10^{-5}) and 0.78 between SLA-7 and SLA-8 (P -value = 10^{-5}). For SLA-Ia, the lowest transcription levels were detected in the cornea, followed by the aorta, brain, skin, kidney and tonsils. Transcription levels two to five times higher than in the brain were detected in lymphoid organs such as the inguinal lymph node, Peyer's patches, spleen and in the organs of the digestive tract such as the duodenum, the jejunum and the ileum. The highest transcription level was found in the lung, but the wide standard deviation value indicated that this high expression is very variable among individual animals. For SLA-6, the highest transcription levels were found in four organs of the lymphoid system, the inguinal lymph node, the thymus, the spleen and Peyer's patches, followed by the lung, organs of the digestive tract and the kidney. The lowest transcription was detected in the skin followed by the brain, the longissimus dorsi, the tonsils and the liver. Transcription of SLA-7 was highest in the lung followed by lymphoid

Table 2 Comparison of the transcription levels of SLA-Ia, SLA-7 and SLA-8 genes to the SLA-6 gene by qRT-PCR.

5

Tissue function/ cell category	Tissues/conditions	Symbols	SLA-Ia ¹		SLA-7 ¹		SLA8 ¹	
			MeLiM	LW	MeLiM	LW	MeLiM	LW
Adult tissues								
Physical barrier	Cornea	C	272.9 (276.3)	—	2.8 (1.2)	—	18.7 (12.9)	—
	Nasal mucosa	NM	537.9 (184.4)	2437.5	3.7 (1.9)	6.8	25.6 (10.4)	41.5
	Skin	Sk	1161.6 (397.4)	830.4	18.6 (15.6)	6.7	46.9 (14.6)	16.4
Fat tissue	Back fat	BF	799.1 (50.5)	714.2	10.9 (11.6)	5.2	52.4 (8.6)	11.3
Skeletal muscle	Diaphragma	Di	755.7 (342.8)	412.4	4 (2.9)	3.1	26.3 (12.5)	12.2
	Longissimus dorsi	LD	761.4 (176.4)	404	5.6 (6.3)	5.4	16.9 (2.1)	11
Respiratory tract	Lung	Lu	1090.3 (1473.6)	197.5	2.6 (1.7)	2.7	14.1 (3.8)	6.5
Cardio-vascular system	Aorta	A	150.8 (38.3)	184	2.5 (2.2)	4.5	10.5 (3.3)	14.8
	Heart	H	338.1 (71.9)	492.1	4 (2.7)	4	14.6 (2.7)	10.6
Nervous system	Brain	B	607.5 (276.1)	1164.9	4.7 (2.9)	6.6	33.9 (10.7)	33.2
Lymphoid system	Inguinal lymph node	IN	168.4 (11.4)	370.4	1.5 (1.1)	1.7	11.2 (3.1)	3.8
	Peyer's patches	PP	520.3 (211.3)	1023.6	2.8 (2.1)	4.6	14.1 (5.2)	11.1
	Spleen	Sp	339.9 (181.1)	88.6	2.1 (1.2)	4.2	18.4 (7.9)	2.5
	Tonsil	T	458.8 (60.6)	283.1	6.7 (5.3)	1	25.5 (11.2)	3.2
Digestive tract and hepatic systems	Thymus	Th	97.7 (61.5)	31.5	1.4 (0.8)	2.6	4.8 (2.4)	0.8
	Duodenum	Du	611.6 (422.6)	500	4.6 (3.2)	3.9	19.2 (13.2)	9.1
	Ileum	I	341.1 (78.4)	295.1	3.2 (2.6)	1.7	14 (4.9)	7.9
	Jejunum	J	452 (194.1)	486.6	3.1 (1.7)	3.3	12.9 (4.1)	15.6
Endocrine gland	Liver	Li	1438.8 (1331.9)	1338.4	8.5 (5)	3.2	28.9 (8.2)	15.4
	Adrenal gland	AG	250.1 (96.8)	1448.1	2.3 (2.1)	8.4	12.4 (3.3)	6.4
Urinary system	Kidney	K	74.6 (61.3)	159.2	2.2 (2)	2.2	7.8 (0.7)	6.6
Genital system	Epididymis	Ep	91.3 (0)	776	2.8 (0)	6.4	5.9 (0)	18
	Ovary	O	264.9 (109)	653	3.3 (3.5)	6	42 (9.8)	19.3
	Testis	Te	1453.3 (0)	634.7	18 (0)	5.6	89.8 (0)	23.9
	Uterus	U	226.8 (118.6)	250.7	2.5 (1.9)	5.9	15 (4.9)	9.6
Average fold change			530	632	5	4	23	12
Cells								
P-K15 epithelial cells	Mock-infected	MI_T0	14.1 (6.7)		2 (2.5)		5.6 (2)	
	T0 post-infection with PrV	I_T0	12.2 (6.2)		2.2 (2.6)		7.4 (3.2)	
	2 h post-infection with PrV	I_T2	12.6 (6.4)		2.3 (2.9)		5.4 (2.8)	
	4 h post-infection with PrV	I_T4	15.3 (9.7)		3 (3.6)		4 (1.5)	
	8 h post-infection with PrV	I_T8	8.2 (4.2)		2.7 (3.4)		1.5 (0.5)	
	12 h post-infection with PrV	I_T12	7.2 (3.5)		1.8 (2)		2.5 (2.9)	
	Average fold change		11		2		4	
PBMCs	Unstimulated	Control	970 (346.5)		0.8 (0.2)		21.5 (8.4)	
	LPS stimulated	LPS	864.2 (401.2)		0.8 (0.6)		21 (11.5)	
	PMA/ionomycin stimulated	PMA/ionomycin	950.3 (243)		1.1 (0.8)		24 (9.4)	
	Average fold change		928		~1		22	

Relative quantification was calculated using SLA-6 data as calibrator. The table summarizes data obtained from various tissues of MeLiM and Large White pigs, from stimulated or non-stimulated PBMCs and from PK15 cells during the time course of infection by PrV.

LPS, lipopolysaccharide; PBMCs, peripheral blood mononuclear cells; PMA, phorbol myristate acetate.

¹The values correspond to the mean fold change and the standard deviations in parenthesis were calculated.

organs (inguinal lymph node, thymus, spleen, Peyer's patches), back fat and organs of the digestive tract. The lowest SLA-7 transcription level was found in the brain. Transcription of SLA-8 was highest in the spleen and the inguinal lymph node followed by the lung, the ovary and Peyer's patches and a group of five tissues including the back fat, the diaphragma, the jejunum, the ileum and the adrenal gland. The lowest transcription was found in the longissimus dorsi and the brain.

Variation in *in vitro* transcription levels of SLA-Ia, SLA-6, SLA-7 and SLA-8 in PBMCs and PK15 cells according to stimulation and infection conditions

During the time course of infection by PrV in PK15 cells (Fig. 4), a slight but significant down-regulation of SLA-Ia genes was found at 8 and 12 h post-infection. A significant decrease in transcription was detected for SLA-8 at 4 and 8 h post-infection, indicating that a detectable down-regu-

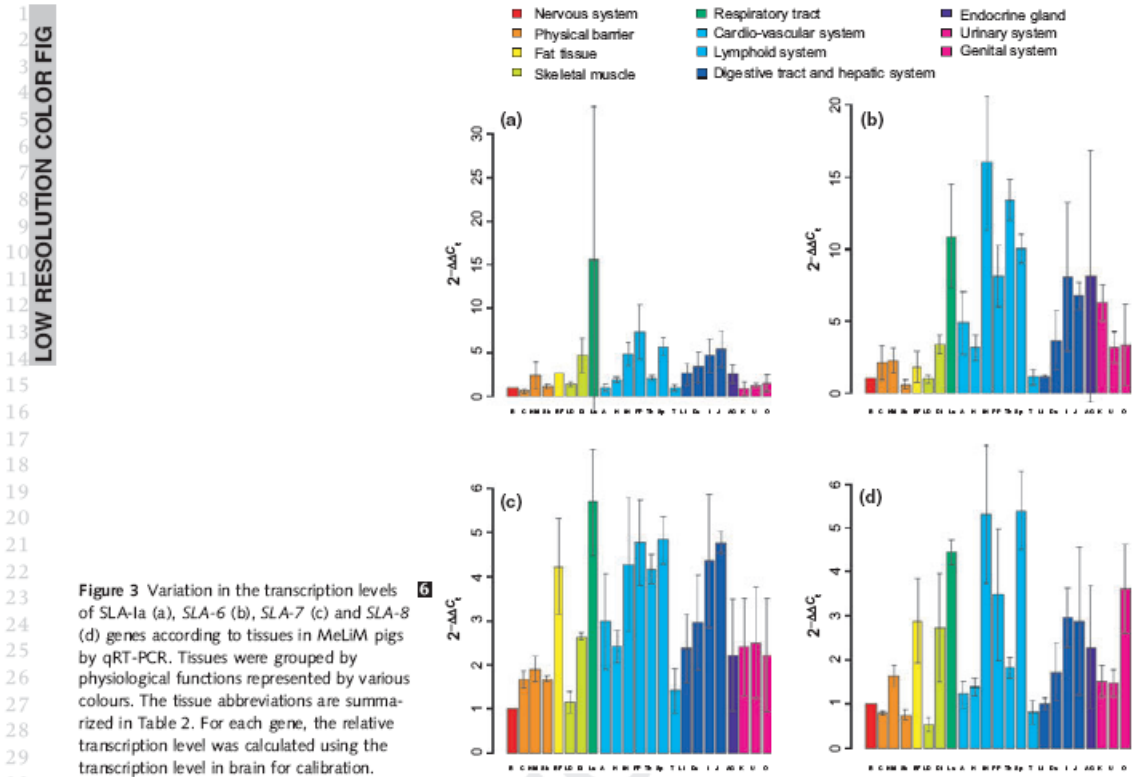


Figure 3 Variation in the transcription levels of SLA-Ia (a), SLA-6 (b), SLA-7 (c) and SLA-8 (d) genes according to tissues in MeLiM pigs by qRT-PCR. Tissues were grouped by physiological functions represented by various colours. The tissue abbreviations are summarized in Table 2. For each gene, the relative transcription level was calculated using the transcription level in brain for calibration.

lation of *SLA-8* occurred earlier than for *SLA-Ia*. No significant variations were identified for *SLA-6* and *SLA-7*, despite a tendency towards down-regulation for *SLA-6* and a tendency towards up-regulation or steady state for *SLA-7*. In PBMCs (Fig. 5), a global down-regulation of classical class I genes was observed after both PMA/ionomycin or LPS stimulation. Similarly, *SLA-8* was found to be down-regulated after both stimulations. In contrast, *SLA-6* was significantly down-regulated after PMA/ionomycin stimulation but up-regulated after LPS stimulation. For *SLA-7*, the variations were not detected as significant.

Discussion

The relative transcription levels of SLA-Ia and SLA-Ib genes were measured by qRT-PCR in 25 adult tissues from two pig breeds as well as in stimulated and unstimulated PBMCs and in infected and non-infected PK-15 cells. They were detected using primers specific for each of the *SLA-6*, *SLA-7* and *SLA-8* genes and a unique primer pair simultaneously targeting the three SLA-Ia genes *SLA-1*, *SLA-2* and *SLA-3*. These primers were shown to work efficiently for both non-related breeds and for the PK-15 cells, indicating that they should be suitable for the transcription studies of various

SLA haplotypes. It is noteworthy that the three classical genes are confounded in our study and are considered as co-expressed for a similar function. We are aware that we cannot rule out the possibility that *SLA-1*, *SLA-2* and *SLA-3* may be expressed at various levels according to haplotypes (Renard *et al.*, personal communication; Tanaka-Matsuda *et al.* 2009). Our results clearly show a significant overexpression of SLA-Ia genes compared to SLA-Ib genes. These findings are in agreement with the classification of the *SLA-6*, *SLA-7* and *SLA-8* genes as non-classical genes, because HLA class Ib genes have also been reported to be less expressed than their classical counterparts (Lury *et al.* 1990; Onno *et al.* 1994; Wainwright *et al.* 2000). Despite significant differences in the average levels of transcription between SLA-Ia and SLA-Ib genes, moderate to high correlations between SLA-Ia and SLA-Ib transcription levels were found.

Our data confirm earlier results on the widespread tissue distribution of SLA-Ib gene expression (Crew *et al.* 2004), but are based on a robust and detailed expression analysis. In humans, *HLA-G* is transcribed in a broad range of tissues (Onno *et al.* 1994), but proteins are mainly detected in the placenta, thymus, cornea, nail matrix, dendritic and allo-reactive T cells (Ishitani *et al.* 2003; Le Discorde *et al.* 2003;

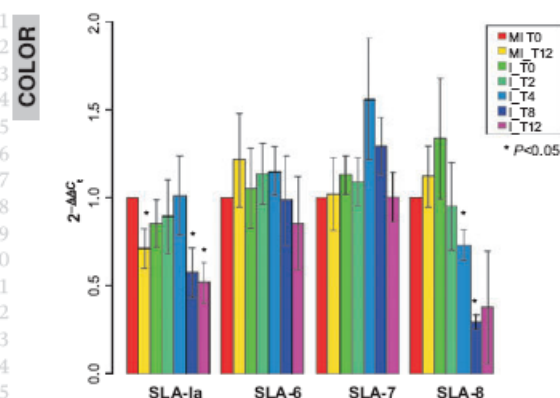


Figure 4 Variation in the transcription level of SLA-Ia, SLA-6, SLA-7 and SLA-8 genes by qRT-PCR in PK15 cells during the time course of infection by PrV. Two conditions of mock infection, just after mock infection (MI_T0) and 12 h after mock infection (MI_T12), are compared with five conditions of PrV infection: just after infection (I_T0), 2 h (I_T2), 4 h (I_T4), 8 h (I_T8) and 12 h (I_T12) after infection. For each condition, three biological replicates have been analysed. Relative quantification was calculated using the transcription level of each gene in mock-infected cells at time 0 post-infection (MI_T0) as a calibrator.

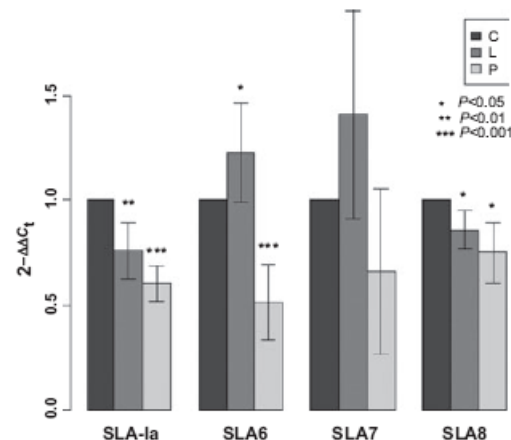


Figure 5 Variation in the transcription level of SLA-Ia, SLA-6, SLA-7 and SLA-8 genes by qRT-PCR in peripheral blood mononuclear cells (PBMCs) either mock-stimulated (C) or stimulated by lipopolysaccharide (L) or phorbol myristate acetate/ionomycin (P) for 24 h. For each condition, seven biological replicates have been analysed. Relative quantification was calculated using the transcription level of each gene in mock-stimulated PBMCs as calibrators.

Le Friec *et al.* 2004; Le Rond *et al.* 2004; Ito *et al.* 2005; Carosella *et al.* 2008). Similarly, *HLA-E* transcripts are found in various cell types (Ulbrecht *et al.* 1992), although cell-surface expression of the proteins seems to be limited to some cell types and/or cell activation processes. We cannot

draw conclusions about SLA-Ib protein expression, owing to the lack of available antibodies that recognize the SLA-Ib molecules. However, a strong correlation between *HLA-G* transcripts and proteins has been reported (Carosella *et al.* 2003), suggesting that the transcription levels of SLA-Ib genes may be a good prediction of the corresponding protein levels. Studies are in progress to produce antibodies specifically recognizing SLA-Ib molecules to study gene expression at the protein level. Our ongoing experiments will provide data to refine correlations between transcript and protein levels.

Our results indicate that *SLA-8* gene is the most transcribed gene, followed by *SLA-7* and finally *SLA-6*, suggesting a possible prominent role of *SLA-8* in adult tissues in basal conditions. However, *SLA-6* harbours the widest range of relative transcription levels according to tissues, suggesting a more specific spectrum of activity. The SLA-Ib genes were weakly transcribed in the brain as previously reported (Crew *et al.* 2004) and as generally observed for MHC genes. The SLA-Ib genes were mainly transcribed in lymphoid tissues except tonsils followed by the lung and organs of the digestive tract. Among the lymphoid organs, *SLA-6* was predominantly expressed in the inguinal lymph node and thymus, *SLA-7* in Peyer's patches and the spleen and *SLA-8* in the inguinal lymph node and spleen. Tonsils are the only lymphoid organs that displayed a low level of SLA-Ib transcription. Notably, a low *HLA-F* protein level has also been detected in adult tonsils (Wainwright *et al.* 2000). Our results show that *SLA-6* is the least transcribed gene in mucosa-associated lymphoid tissues (tonsils and Peyer's patches) and *SLA-8* is the most predominantly transcribed gene in secondary organs (spleen and lymph node), suggesting a more specific expression of *SLA-6* and *SLA-8* in lymphoid organs than *SLA-7*. Initially, *SLA-6* has been referred to as *PD6* and its expression was shown to be mostly restricted to secondary lymphoid tissues such as spleen and lymph node (Ehrlich *et al.* 1987). However, our results show that its expression level in thymus is intermediate between that in lymph node and spleen. *SLA-7* is the most transcribed gene in lung, and *SLA-6* and *SLA-8* are also highly transcribed in this organ. Interestingly, transcription of SLA-Ib genes was also high in the jejunum, ileum and duodenum, providing the third position after lymphoid organs and lung to the digestive tract in biological functions. The role of non-classical genes in the cross-talk between microbiota and intestine epithelium is an emerging field, and it is not excluded that the SLA-Ib genes may have a non-negligible role in pig. Other non-classical genes located outside the MHC locus such as *CD1d* (Russano *et al.* 2007; Salio *et al.* 2010) and *MR1* (Huang *et al.* 2008) are reported to specifically present peptides in this micro-environment (Middendorp & Nieuwenhuis 2009).

We have previously shown that the SLA-Ia genes are down-regulated after PMA/ionomycin stimulation (Gao *et al.* 2010) and have also confirmed other studies on the

down-regulation of SLA-Ia genes in PK15 cells infected by the PrV (Mellencamp *et al.* 1991; Flori *et al.* 2008a,b). The SLA-Ia and SLA-8 genes were coregulated in PK-15 cells during the time course of infection or in PBMCs according to stimulation in contrast to SLA-6 and SLA-7 that were variously modulated. Indeed, a down-regulation of SLA-8 was observed at 4 and 8 h post-infection by the PrV infection and 24 h after LPS or PMA/ionomycin stimulations. Analysis of the promoter activity of SLA-Ia and SLA-6 and SLA-7 genes by *in vitro* transfection of expression vectors and co-expression of interferon regulatory factors or cell treatment with TNF- α have shown distinct regulations of the genes (Tennant *et al.* 2007). Combined with our results, these findings are in agreement with distinct modulations of expression of SLA-6 and SLA-7 compared to SLA-Ia after immune stimulation.

For adult tissues, we included animals from two pig breeds. French Large White pigs correspond to commercial animals highly selected for meat production and quality and bred in an experimental farm unit. MeLiM animals belong to an inbred closed line selected for the spontaneous occurrence and regression of cutaneous melanomas (Vincent-Naulleau *et al.* 2004) and are maintained in our experimental facilities for biomedical research. Healthy MeLiM animals have been selected for our study. Our results provide insights into the physiological distribution of the SLA-Ia and SLA-Ib gene transcripts in tissues derived from five distinct individuals. If the three Melim animals provided more robust and significant results than the only two Large White animals included in our study, the relative expression of the genes per tissue is in the same order of magnitude for most tissues in both breeds. Two distinct internal reference genes have been used for the qRT-PCRs but the comparable orders of magnitude obtained with both species suggest that using B2M or RPL32 as a reference gene has no major impact on the results. The discrepancies observed between breeds for some tissues could be attributed to the breeding environment, which was different for the two breeds and might significantly affect the animal's immune status even if all animals were considered to be healthy. We agree that as breeds and environments are confounded in our experiment, it would be premature to conclude that results on tissue specificity are constant whatever the breed, individual or environment. However, the interesting fact is that the average relative ratios between the gene transcription levels are maintained for both breeds, suggesting a constant predominant expression of SLA-8 in comparison with SLA-6 and SLA-7. This result does not preclude a modification in this relative order of gene expression in pathological or stress conditions not yet studied. It is noteworthy that the relative expression of classical class Ia genes compared to Ib genes was found to be 10–100 times lower in PK15 cells than in *in vivo* (adult tissues) or *ex vivo* (stimulated and non-stimulated PBMCs) conditions.

A clear correspondence between SLA-Ib and HLA-Ib genes based on the gene expression levels in different tissues is difficult to establish based on our results. However, as SLA-8 is the most SLA-Ib transcribed gene in the tissues and conditions included in our study, it is tempting to hypothesize that it may be a functional homologue to HLA-E, which is reported to be expressed in a wide range of cell types (Lee *et al.* 1998). The role of HLA-E in presenting peptides derived from MHC class I leader sequences and in binding to various immunoreceptors is shared by higher primates and mouse and might be common to more mammalian species including the pig. Sequence orthologies are clear in primates (Knapp *et al.* 1998), and a functional ortholog has been identified as H2-Qa1 in mouse (Kurepa *et al.* 1998; Joly & Rouillon 2006).

A specific study of the SLA-Ib gene expression level in various tissues of the female reproductive tract and at various positions of the foeto-maternal interface during the time course of gestation would add complementary data to the present study. Experiments are planned for this specific purpose. Placentation is hemochorial in the human species whereas it is non-invasive and epitheliochorial in the pig species, suggesting that the role of immunotolerant molecules such as HLA-Ib molecules in the pig species during gestation could be less crucial than in humans. The tissue-wide transcription of the three SLA-Ib genes in various adult tissues does not preclude an active role during gestation, but suggests additional functions that might differ from known functions in humans such as for induction and maintenance of a foeto-maternal tolerance during pregnancy. Functional experiments are in progress to study the ability of SLA-6, SLA-7 or SLA-8 molecules to bind B2M and migrate to the cell surface. These results will help to elucidate SLA-Ib gene roles and identify putative functional homologies with MHC class Ib genes of other species. Much work is still required to assess whether these genes encode molecules that have roles shared by other species and/or that fulfil functions not yet known and specific to the pig species.

Acknowledgements

The authors thank Steliana Kevorkian (Multiple Users Research Center-Molecular Biology, University of Bucharest), who efficiently worked on the project during her master degree. The authors are grateful to Claudia Bevilacqua (INRA-GABI, ICE platform) and Jean-Pierre Furet (INRA-MICALIS) for their help in qRT-PCR experiments, as well as Jean-Jacques Leplat (INRA-GABI and CEA-DSV-IRCM-SREIT-LREG) and Stephan Bouet (INRA-GABI) for their contribution to tissue sampling on MeLiM pigs. The authors also thank Hélène Hayes (INRA-GABI) for correcting the manuscript and Bertrand Bed'hom (INRA-GABI) for helpful suggestions. This research project has been supported by a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework

Animal Genetics © 2010 Stichting International Foundation for Animal Genetics

No claim to original US government works, doi: 10.1111/j.1365-2052.2010.02170.x

Programme under contract number MEST-CT-2004-504854 and by the Animal Genetics Department of INRA.

References

- Anfossi N., Andre P., Guia S. et al. (2006) Human NK cell education by inhibitory receptors for MHC class I. *Immunity* 25, 331–42.
- Apps R., Gardner L., Traherne J., Male V. & Moffett A. (2008) Natural-killer cell ligands at the maternal-fetal interface: UL-16 binding proteins, MHC class-I chain related molecules, HLA-F and CD48. *Human Reproduction* 23, 2535–48.
- Birch J., Codner G., Guzman E. & Ellis S.A. (2008) Genomic location and characterisation of nonclassical MHC class I genes in cattle. *Immunogenetics* 60, 267–73.
- Boyle L.H., Gillingham A.K., Munro S. & Trowsdale J. (2006) Selective export of HLA-F by its cytoplasmic tail. *Journal of Immunology* 176, 6464–72.
- Carosella E.D., Moreau P., Le Maout J., Le Discorde M., Dausset J. & Rouas-Freiss N. (2003) HLA-G molecules: from maternal-fetal tolerance to tissue acceptance. *Advances in Immunology* 81, 199–252.
- Carosella E.D., Favier B., Rouas-Freiss N., Moreau P. & Lemaout J. (2008) Beyond the increasing complexity of the immunomodulatory HLA-G molecule. *Blood* 111, 4862–70.
- Chardon P., Rogel-Gaillard C., Cattolico L., Duprat S., Vaiman M. & Renard C. (2001) Sequence of the swine major histocompatibility complex region containing all non-classical class I genes. *Tissue Antigens* 57, 55–65.
- Comiskey M., Goldstein C.Y., De Fazio S.R., Mammolenti M., Newmark J.A. & Warner C.M. (2003) Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product. *Human Immunology* 64, 999–1004.
- Crew M.D., Phanavanh B. & Garcia-Borges C.N. (2004) Sequence and mRNA expression of nonclassical SLA class I genes SLA-7 and SLA-8. *Immunogenetics* 56, 111–4.
- Donnelly C.E., Yatko C., Johnson E.W. & Edge A.S. (1997) Human natural killer cells account for non-MHC class I-restricted cytotoxicity of porcine cells. *Cellular Immunology* 175, 171–8.
- Ehrlich R., Lifshitz R., Pescovitz M.D., Rudikoff S. & Singer D.S. (1987) Tissue-specific expression and structure of a divergent member of a class I MHC gene family. *Journal of Immunology* 139, 593–602.
- Flori L., Rogel-Gaillard C., Cochet M., Lemonnier G., Hugot K., Chardon P., Robin S. & Lefevre F. (2008a) Transcriptomic analysis of the dialogue between Pseudorabies virus and porcine epithelial cells during infection. *BMC Genomics* 9, 123.
- Flori L., Rogel-Gaillard C., Mariani V., Lemonnier G., Cochet M., Hugot K., Chardon P., Robin S. & Lefevre F. (2008b) A combined transcriptomic approach to analyse the dialogue between pseudorabies virus and porcine cells. *Developments in Biologicals* 132, 99–104.
- Gao Y., Flori L., Lecardonnel J., Esquerre D., Hu Z.L., Teillaud A., Lemonnier G., Lefevre F., Oswald I.P. & Rogel-Gaillard C. (2010) Transcriptome analysis of porcine PBMCs after in vitro stimulation by IPS or PMA/ionomycin using an expression array targeting the pig immune response. *BMC Genomics* 11, 292.
- Gomes A.Q., Correia D.V. & Silva-Santos B. (2007) Non-classical major histocompatibility complex proteins as determinants of tumour immunosurveillance. *EMBO Reports* 8, 1024–30.
- Ho C.S., Lunney J.K., Ando A., Rogel-Gaillard C., Lee J.H., Schook L.B. & Smith D.M. (2009) Nomenclature for factors of the SLA system, update 2008. *Tissue Antigens* 73, 307–15.
- Horak V., Fortyn K., Hruban V. & Klaudy J. (1999) Hereditary melanoblastoma in miniature pigs and its successful therapy by devitalization technique. *Cellular and Molecular Biology (Noisy-le-Grand, France)* 45, 1119–29.
- Horton R., Wilming L., Rand V. et al. (2004) Gene map of the extended human MHC. *Nature Reviews. Genetics* 5, 889–99.
- Huang S., Gilfillan S., Kim S., Thompson B., Wang X., Sant A.J., Fremont D.H., Lantz O. & Hansen T.H. (2008) MR1 uses an endocytic pathway to activate mucosal-associated invariant T cells. *The Journal of Experimental Medicine* 205, 1201–11.
- Hunt J.S., Morales P.J., Pace J.L., Fazleabas A.T. & Langat D.K. (2007) A commentary on gestational programming and functions of HLA-G in pregnancy. *Placenta* 28(Suppl. A), S57–63.
- Ishitani A., Sageshima N., Lee N., Dorofeeva N., Hatake K., Marquardt H. & Geraghty D.E. (2003) Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. *Journal of Immunology* 171, 1376–84.
- Itescu S., Kwiatkowski P., Artrip J.H., Wang S.F., Ankersmit J., Minanov O.P. & Michler R.E. (1998) Role of natural killer cells, macrophages, and accessory molecule interactions in the rejection of pig-to-primate xenografts beyond the hyperacute period. *Human Immunology* 59, 275–86.
- Ito T., Ito N., Saathoff M., Stampachiachiere B., Bettermann A., Bullone-Paus S., Takigawa M., Nickoloff B.J. & Paus R. (2005) Immunology of the human nail apparatus: the nail matrix is a site of relative immune privilege. *The Journal of Investigative Dermatology* 125, 1139–48.
- Joly E. & Rouillon V. (2006) The orthology of HLA-E and H2-Qa1 is hidden by their concerted evolution with other MHC class I molecules. *Biology Direct* 1, 2.
- Kelley J., Walter L. & Trowsdale J. (2005) Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56, 683–95.
- Knapp L.A., Cadavid L.F. & Watkins D.I. (1998) The MHC-E locus is the most well conserved of all known primate class I histocompatibility genes. *Journal of Immunology* 160, 189–96.
- Kurepa Z., Hasemann C.A. & Forman J. (1998) Qa-1b binds conserved class I leader peptides derived from several mammalian species. *The Journal of Experimental Medicine* 188, 973–8.
- Lanier L.L. (2001) On guard-activating NK cell receptors. *Nature Immunology* 2, 23–7.
- Larkin M.A., Blackshields G., Brown N.P. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–8.
- Le Discorde M., Moreau P., Sabatier P., Legeais J.M. & Carosella E.D. (2003) Expression of HLA-G in human cornea, an immune-privileged tissue. *Human Immunology* 64, 1039–44.
- Le Fric G., Gros F., Sebt Y., Guilloux V., Pangault C., Fauchet R. & Amiot L. (2004) Capacity of myeloid and plasmacytoid dendritic cells especially at mature stage to express and secrete HLA-G molecules. *Journal of Leukocyte Biology* 76, 1125–33.
- Le Rond S., Le Maout J., Creput C. et al. (2004) Alloreactive CD4+ and CD8+ T cells express the immunotolerant HLA-G molecule in mixed lymphocyte reactions: in vivo implications in transplanted patients. *European Journal of Immunology* 34, 649–60.

- Lee N., Llano M., Carretero M., Ishitani A., Navarro F., Lopez-Botet M. & Geraghty D.E. (1998) HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. *Proceedings of the National Academy of Sciences of the United States of America* 95, 5199–204.
- LeMaoult J., Rouas-Freiss N. & Carosella E.D. (2005) Immunotolerogenic functions of HLA-G: relevance in transplantation and oncology. *Autoimmunity Reviews* 4, 503–9.
- Lunney J.K., Ho C.S., Wysocki M. & Smith D.M. (2009) Molecular genetics of the swine major histocompatibility complex, the SLA complex. *Developmental and Comparative Immunology* 33, 362–74.
- Lury D., Epstein H. & Holmes N. (1990) The human class I MHC gene HLA-F is expressed in lymphocytes. *International Immunology* 2, 531–7.
- Mellencamp M.W., O'Brien P.C. & Stevenson J.R. (1991) Pseudorabies virus-induced suppression of major histocompatibility complex class I antigen expression. *Journal of Virology* 65, 3365–8.
- Middendorp S. & Nieuwenhuis E.E. (2009) NKT cells in mucosal immunity. *Mucosal Immunology* 2, 393–402.
- Moscoso J., Serrano-Vela J.L., Pacheco R. & Arnaiz-Villena A. (2006) HLA-G, -E and -F: allelism, function and evolution. *Transplant Immunology* 17, 61–4.
- Onno M., Guillaudeux T., Amiot L., Renard I., Drenou B., Hirel B., Girr M., Semana G., Le Bouteiller P. & Fauchet R. (1994) The HLA-G gene is expressed at a low mRNA level in different human cells and tissues. *Human Immunology* 41, 79–86.
- Renard C., Hart E., Sehra H. *et al.* (2006) The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics* 88, 96–110.
- Rock K.L., Farfan-Arribas D.J. & Shen L. (2010) Proteases in MHC class I presentation and cross-presentation. *Journal of Immunology* 184, 9–15.
- Rodgers J.R. & Cook R.G. (2005) MHC class Ib molecules bridge innate and acquired immunity. *Nature Reviews Immunology* 5, 459–71.
- Russano A.M., Bassotti G., Agea E., Bistoni O., Mazzocchi A., Morelli A., Porcelli S.A. & Spinozzi F. (2007) CD1-restricted recognition of exogenous and self-lipid antigens by duodenal gammadelta+ T lymphocytes. *Journal of Immunology* 178, 3620–6.
- Salio M., Silk J.D. & Cerundolo V. (2010) Recent advances in processing and presentation of CD1 bound lipid antigens. *Current Opinion in Immunology* 22, 81–8.
- Shawar S.M., Vyas J.M., Rodgers J.R. & Rich R.R. (1994) Antigen presentation by major histocompatibility complex class I-B molecules. *Annual Review of Immunology* 12, 839–80.
- Shishido S., Naziruddin B., Howard T. & Mohanakumar T. (1997) Recognition of porcine major histocompatibility complex class I antigens by human CD8+ cytolytic T cell clones. *Transplantation* 64, 340–6.
- Smith D.M., Lunney J.K., Martens G.W., Ando A., Lee J.H., Ho C.S., Schook L., Renard C. & Chardon P. (2005) Nomenclature for factors of the SLA class-I system, 2004. *Tissue Antigens* 65, 136–49.
- Sullivan L.C., Clements C.S., Rossjohn J. & Brooks A.G. (2008) The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. *Tissue Antigens* 72, 222–222.
- Tanaka-Matsuda M., Ando A., Rogel-Gaillard C., Chardon P. & Uenishi H. (2009) Difference in number of loci of swine leukocyte antigen classical class I genes among haplotypes. *Genomics* 93, 261–73.
- Tennant L.M., Renard C., Chardon P. & Powell P.P. (2007) Regulation of porcine classical and nonclassical MHC class I expression. *Immunogenetics* 59, 222–222.
- Ulbrecht M., Kellermann J., Johnson J.P. & Weiss E.H. (1992) Impaired intracellular transport and cell surface expression of nonpolymorphic HLA-E: evidence for inefficient peptide binding. *The Journal of Experimental Medicine* 176, 1083–90.
- Vincent-Naulleau S., Le Chalony C., Leplat J.J. *et al.* (2004) Clinical and histopathological characterization of cutaneous melanomas in the melanoblastoma-bearing Libechev minipig model. *Pigment Cell Research* 17, 24–35.
- Wainwright S.D., Biro P.A. & Holmes C.H. (2000) HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. *Journal of Immunology* 164, 319–28.
- Xu X.C., Naziruddin B., Sasaki H., Smith D.M. & Mohanakumar T. (1999) Allele-specific and peptide-dependent recognition of swine leukocyte antigen class I by human cytotoxic T-cell clones. *Transplantation* 68, 473–9.

Abstracts and posters

International Symposium on Animal Genomics for Animal Health, May 31 – June 2, 2010: poster

Transcription variants of the non classical MHC class I genes in pig

Rui HU^{1,2}, Gaëtan LEMONNIER^{1,2}, Emmanuelle BOURNEUF^{1,2}, Silvia VINCENT-NAULLEAU^{1,2}, Claire ROGEL-GAILLARD^{1,2}

¹INRA, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

²CEA, DSV, IRCM, Laboratoire de Radiobiologie et Etude du Génome, Jouy-en-Josas, France

In pig, very little information is available on the three non classical MHC class I genes SLA-6, -7 and -8 (SLA-Ib genes). In order to study the transcription pattern of these three genes, RT-PCRs with gene specific primers were carried out. Full length transcripts were characterized from thymus and brain of MeLiM pigs resulting in the annotation of 8 exons for SLA-7 and -8 and 7 exons for SLA-6. The three full length cDNAs encode molecules with a predicted folding consistent with peptide presentation. No additional transcript was found for SLA-8 while four were detected for SLA-6. One variant encodes a putative truncated protein because intron 3 is retained and creates a premature termination codon. Three alternatively spliced variants encode putative proteins lacking exon 3 (alpha2 domain), exons 3 and 5 (transmembrane domain) or exons 3, 5 and 6 (cytoplasmic tail) suggesting that soluble forms of SLA-6 molecules may exist. For SLA-7, an alternative spliced variant was found in the 3'UTR of the gene after the termination codon suggesting possible post-transcriptional regulation of the gene. In conclusion, our results show that SLA-6 and -7 genes express alternatively spliced transcripts in contrast to SLA-8. These transcription patterns were confirmed in 23 distinct tissues of Large White and MeLiM tissues. Since no orthology between SLA-Ib and HLA-Ib genes known as HLA-E, -F and -G has been established at the sequence level, functional homologies will be discussed on the basis of protein prediction.

Transcription variants of the non classical MHC class I genes in pig

Rui HU^{1,2}, Gaëtan LEMONNIER^{1,2}, Emmanuelle BOURNEUF^{1,2}, Silvia VINCENT-NAULLEAU^{1,2}, Claire ROGEL-GAILLARD^{1,2}

¹ INRA, AgroParistech, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

² CEA, DSV, IRCM, SREIT, Laboratoire de Radiobiologie et Etude du Génome, Jouy-en-Josas, France

Three non classical MHC class I (class Ib) genes have been identified in the pig, namely SLA-6, -7 and -8. Very little information is available on these three genes. MHC class Ib genes belong to the MHC class I gene family but differ from classical genes by a more restricted tissue specificity, a low polymorphism, possible alternative transcription and alterations in the three prime end region encoding the cytoplasmic tail. In human, MHC class Ib genes were shown to have a role in foeto-maternal tolerance during pregnancy¹ and immunotolerance in tumors^{1,2}. As a prerequisite toward functional studies on SLA-Ib genes, our objective was to characterize their transcriptional pattern in two pig breeds.

Materials and methods

- RT-PCRs with gene specific primers were carried out. Full length transcripts were characterized from thymus and brain of MeLiM pigs using primers from exon 1 to 3'UTR (Fig. 1).

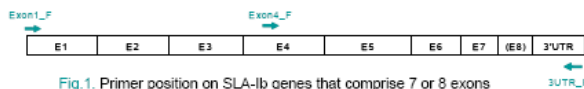


Fig.1. Primer position on SLA-Ib genes that comprise 7 or 8 exons

- 24 tissues from Large White and MeLiM animals (Fig.2) were screened for SLA-Ib transcription by RT-PCR with primers targeting the transcripts from exon 4 to the 3'end (Fig.1).



Fig.2. Pig breeds included in the study.
A. MeLiM pig: Melanoma bearing Lieberov Minipig³ (biomedical model)
B. French Large White pig (selection for production traits)

- The protein structure prediction was done with CPHmodels 3.0 (<http://www.cbs.dtu.dk/services/CPHmodels/>), which is a protein homology modeling server based on profile-profile alignment guided by secondary structure and exposure predictions.

Results

Identification of transcription variants for SLA-6 and -7 genes but detection of a unique transcript for SLA-8 gene (Fig.3)

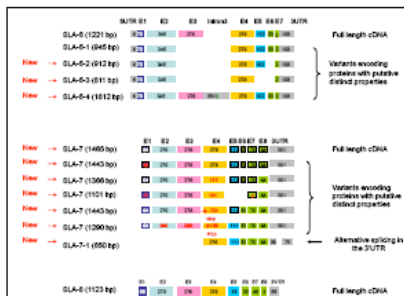


Fig.3. SLA-Ib transcripts found in MeLiM pigs. Arrows point to newly found transcripts.

The SLA-6 and -7 variants of transcription encode putative membrane-bound and soluble proteins (Fig.4). The three SLA-Ib full length cDNAs encode proteins that have the putative capacity to associate with B2M and to bind peptides (Fig. 4 and 5).

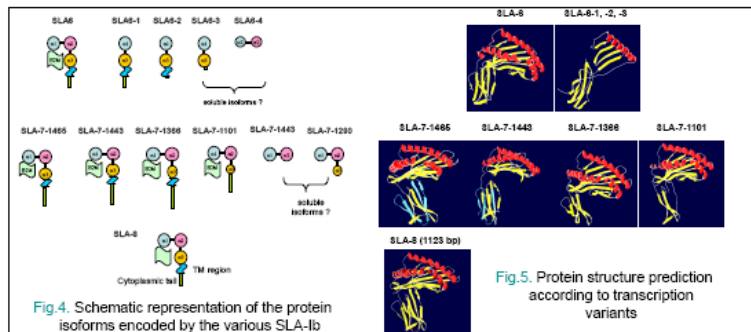


Fig.4. Schematic representation of the protein isoforms encoded by the various SLA-Ib transcripts.

Fig.5. Protein structure prediction according to transcription variants

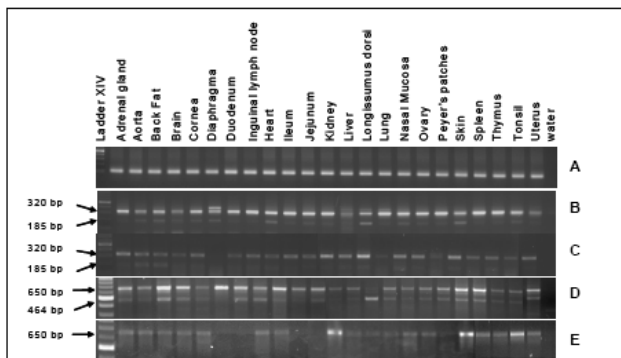


Fig.6. Detection of SLA-6 and -7 transcription variants in tissues of adult MeLiM (B, D) and Large White (C, E) animals. The RT-PCRs target the 3'end of the transcripts (Fig.1). A: RT-PCR on MeLiM tissues with primers specific of the RPL32 gene.

Conclusion

- Alternative transcription for SLA-6 and -7 genes but not for SLA-8 gene
- Variation in the relative transcription levels of the SLA-6 and -7 transcript variants according to tissues
- SLA-7-1 variant predominantly expressed in MeLiM pigs by comparison with Large White pigs
- Ongoing experiments to detect the proteins

References

- Carosella et al. 2003. HLA-G molecules: from maternal-fetal tolerance to tissue acceptance. *Adv Immunol* 81, 199-252.
- Gomes et al. 2007. Non-classical major histocompatibility complex proteins as determinants of tumour immunosurveillance. *EBIO Rep* 6, 104-106.
- Vincent-Naualeau et al. 2004. Clinical and histopathological characterization of cutaneous melanomas in the melanoma-bearing Lieberov minipig model. *Pigment Cell Res* 17, 24-35.
- Lund et al. 2002. CPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract at the CASP5 conference A102, 2002.



32nd Conference of the International Society for Animal Genetics

Edinburgh, 26 – 30 July 2010: Student fellowship, poster and oral presentation

Transcription patterns and polymorphism of the non classical MHC class I genes in pig

Rui HU^{1,2}, Emmanuelle BOURNEUF^{1,2}, Bertrand BED'HOM^{1,2}, Silvia VINCENT-NAULLEAU^{1,2}, Claire ROGEL-GAILLARD^{1,2}

1. INRA, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

2. CEA, DSV, IRCM, SREIT, Laboratoire de Radiobiologie et Etude du Génome, Jouy-en-Josas, France

The major histocompatibility complex (MHC) class I genes comprises classical (Ia) and non classical (Ib) genes. The highly polymorphic class Ia molecules are widely expressed, present peptides to cytotoxic T cells and modulate activity of natural killer cells. In contrast, class Ib genes are predominantly expressed in immunotolerant organ sites, notably at the foeto maternal interface during pregnancy in human, harbour limited polymorphism and can be alternatively spliced. In pig, very limited information is available on the class Ib genes SLA-6, -7 and -8. Our aim was to study transcription patterns and polymorphism of these genes. RT-PCRs with gene specific primers were carried out. Full length transcripts were characterized resulting in annotation of 8 exons for SLA-7 and -8 and 7 exons for SLA-6. For SLA-8, no additional transcript was found. For SLA-6, four additional transcripts were detected and for SLA-7, an alternative spliced variant was found in the 3'UTR of the gene after the termination codon suggesting possible post-transcriptional regulation. For polymorphism studies, long PCRs suitable for amplification of the three genes from the five prime downstream to the three prime non coding sequence were designed and MeLiM pigs harbouring spontaneous melanomas were characterized. Sequencing results confirm the low level of polymorphism of the three genes. In order to start functional studies, ongoing experiments aiming at detecting expression of the proteins at the cell surface will be carried out in a near future. Our overall results will provide very new data on these enigmatic SLA-Ib molecules.

Transcription variants and polymorphism study of non classical MHC class I genes in pig

Rui HU^{1,2}, Emmanuelle BOURNEUF^{1,2}, Bertrand Bed'Hom^{1,2}, Silvia VINCENT-NAULLEAU^{1,2}, Claire ROGEL-GAILLARD^{1,2}

¹ INRA, AgroParisTech, UMR de Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

² CEA, DSV, IRCM, SREIT, Laboratoire de Radiobiologie et Etude du Génome, Jouy-en-Josas, France

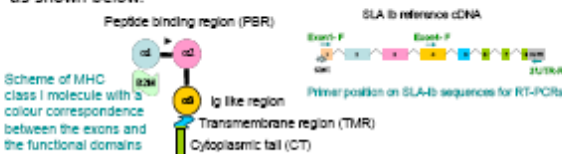
The major histocompatibility complex (MHC) class I genes comprises classical (Ia) and non classical (Ib) genes. The highly polymorphic class Ia molecules are widely expressed, present peptides to cytotoxic T cells and modulate activity of natural killer cells. In contrast, class Ib genes are predominantly expressed in immunotolerant organ sites, notably at the foeto maternal interface during pregnancy in human, harbour limited polymorphism and can be alternatively spliced. In pig, very limited information is available on the class Ib genes SLA-6, -7 and -8. The aim of the project is to study the transcription patterns and polymorphisms of these genes.

Materials and methods

Animals



RT-PCRs with gene specific primers were carried out. Full length and partial transcripts were characterized from thymus and brain of MeLiM pigs using primers derived from exons 1 or 4 and the 3'UTR as shown below.



SLA Ib transcription

SLA-6, -7 and -8 genes annotations were retrieved from VEGA database (http://vega.sanger.ac.uk/Sus_scrofa/info/index) and previously reported full length cDNAs (Crew et al., 2004, Immunogenetics, 56:111) were used as reference data for description of new transcripts.

SLA-6 : 5 transcript variants

- Full length coding sequence: 7 exons
- Three variants with missing exons and one variant with a retained intron



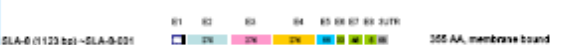
SLA-7 : 6 new transcript variants

- Refinement of the full length coding sequence: 8 exons
- 2 transcripts encoding only the leader peptide
- 1 variant with an alternative splicing in the 3'UTR
- 1 variant using a non canonical GA-AG splice site in exon 4 (splice site rarely used, see Bradley et al., 2005, J Med Genet 2005, 42(8):e51)
- 1 variant with 18 nt missing within exon 4: RNA splicing or genomic deletion ?



SLA-8 : A unique transcript

- Full length coding sequence: 8 exons



Long range PCRs (QIAGEN Long Range PCR kit) were carried out on genomic DNA for further cloning (Invitrogen, pCR-XL-TOPO kit) and sequencing of the SLA-7 gene. The amplified genomic fragment (4145 base pairs) spans the entire SLA-7 gene upstream from the promoter and downstream to the polyadenylation signal.

Sequence analysis and SNP detection were carried out using the CodonCode software (<http://www.codoncode.com/>).

Pyrosequencing was carried out (QIAGEN, Pyromark Q24) to target various SNPs



SLA Ib polymorphism

6 MeLiM pigs were studied for SLA-7 polymorphism and 4 to 6 clones per animal were sequenced.

Animal	MeLiM C284	MeLiM C321	MeLiM B52	MeLiM F206	MeLiM F213	MeLiM 486
# clones	5	4	5	4	4	4
# alleles	3	3	4	2	2	4

An unexpected number of distinct sequences:

- 2 to 4 SLA-7 alleles per animal

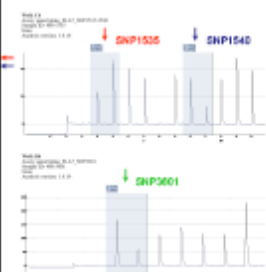
Analysis of animal MeLiM 486 SNP positions:

- exons 2, 3, 4 and 7
- 5' and 3' ends of the gene; introns 1, 2, 5 and 7

Region	Exon	Start	End	SNP	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6
Exon 1	1	284	324	285	A	G	A	A	A	A
				286	A	A	A	A	A	A
				287	A	A	A	A	A	A
				288	A	A	A	A	A	A
Exon 2	2	345	385	346	A	G	A	A	A	A
				347	A	A	A	A	A	A
				348	A	A	A	A	A	A
				349	A	A	A	A	A	A
Exon 3	3	406	446	407	A	G	A	A	A	A
				408	A	A	A	A	A	A
				409	A	A	A	A	A	A
				410	A	A	A	A	A	A
Exon 4	4	467	507	468	A	G	A	A	A	A
				469	A	A	A	A	A	A
				470	A	A	A	A	A	A
				471	A	A	A	A	A	A
Exon 5	5	528	568	529	A	G	A	A	A	A
				530	A	A	A	A	A	A
				531	A	A	A	A	A	A
				532	A	A	A	A	A	A
Exon 6	6	589	629	590	A	G	A	A	A	A
				591	A	A	A	A	A	A
				592	A	A	A	A	A	A
				593	A	A	A	A	A	A
Exon 7	7	650	690	651	A	G	A	A	A	A
				652	A	A	A	A	A	A
				653	A	A	A	A	A	A
				654	A	A	A	A	A	A
Exon 8	8	711	751	712	A	G	A	A	A	A
				713	A	A	A	A	A	A
				714	A	A	A	A	A	A
				715	A	A	A	A	A	A

Pyrosequencing results

- SNP1535: 60% A 40% G
- SNP1540: 73% G 27% T
- SNP3601: 51% C 49% T



Hypothesis:

- SLA-7 gene duplication

Summary

- Alternative transcription for SLA-6 and -7 genes but not for SLA-8
- A transcription pattern more complex for SLA-7 than for SLA-6
- Some putative encoded SLA-6 and SLA-7 protein isoforms may be soluble
- Duplication of SLA-7 gene in MeLiM animals

Supplementary files

Supplementary file S1: Nucleotide sequences of the cDNAs clones into expression vectors from initiation to stop codons with exon delineation for SLA sequences

Supplementary file S2: Translated sequences of the cDNAs cloned into expression vectors with exon delineation for the SLA proteins

Supplementary file S 3: DNA vaccination protocol and specificity of polyclonal serums

Supplementary file S1: Nucleotide sequences of the cDNAs clones into expression vectors
from initiation to stop codons with exon delineation for SLA sequences

> MeLiM485_SLA-6 1110 bp

ATGCAGGTCACGGAGCCTCGAACCCTCCTGCTGGTGCTCTCGGGGTCCCTGGCCCTGACCGAGACCTGGGCAGGCCACCTTC
ACCACGGACCCAAGGACCCGCGTCTGGAGAAGGGTCAGACCCGGCCTCAGCCCCTCCCTGTTATAGGATCCCACCTCGCTAAG
ATACCTCCATATTTTGGTATCCCGGCCCGGCCACGGTAGTGACCTTTACAGCTCTGTGCGTTTCTTGGACGACACGCAGTTC
GTACGGTTTCAGCAGCGATGCCGGAATCCAAGGGTGGAGCCTCGGGCGCGTGGATGGAGCAGGAGGGCGGGAATATTGGG
ATCGTCAGACTGATATAGCCAAAGAACATTCAAGGCTTCAAGATCGAACCTGCGGGTGATCATTGGCAACCACAACCATAG
CCAGTCGGAGTCGCACAGCTTTCTTTGGGTATCCGGCTGCGACGTGGGATCCGACGGGCGCATCCTTCGAGGGTACGAGCAG
TTCTCCTACGACGGCGACGATTACATCGTCCTGAATGAAGACCTGCGCTCCTGGACCGCATCAGTACAGTGGCTCAGATCA
TCCGGCGCAAGTGGGAGGCGGAGGGAGTAGCTGAACAGTATCGGGCATACTTGGAGATCGAGTGTGTGGAGTGGCTCCGCAA
GTACCTCGAGAAGGGGAAGGACGTGCTGCAGCGCGCAGTTCCCTCCAAAGACTCATGTGACCCGCCACCCCTTCTATGACAA
AAAGTCACTCTAAGATGCTGGGCCCTGGGCTTCTACCCTAAGGAGATCTCCCTGACCTGGCAGCGTGATGGGGAGGACCAGA
CCCAGGACATGGAGCTCGTGGAGACCAGGCCTTCAGGGGATGGGACCTTCCAGAAGTGGGCGGCCCTGGTGGTGCCTTCTGG
AGAGGAGCAGAGCTACACCTGCCAGGTGCAGCACGAGGGCCTGCAGGAGCCCCTCACCTGAGATGGGAGCCACCTCCTCTG
TCCCTCAACCTCATCATGATTTGCATTCCTGTGAGTGTGCTTGTGGTCACTGTGCTTGGAACTGTGATCTGGAGAAAGAGGA
ACTCAGGTGGAAATAGAGGGAACATATGTTCAAGCTGCAAGCTGA

> MeLiM485_SLA-7 1167 bp

ATGGGGCCCCAAGCCCTCCTCCTGCTGCTCTCCGGGACCCCTGGTATTGACCCAGCCCTGGGCACGTCCCCACTCCCTGAGGT
ATTTCTACACAGCCGTGTCCCGACCGAGTCGACGGGATCCCGCTTCTCCGTCGTCGGCTATGTAGACGACACGCAGTTCGT
GCGGTTTCGACAGCAATGCCCCGAATCCGAGGGAGGAGCCGCGGACACCGTGGGTGGAGCTGGAGGGCCCGGAGTATTGCGAT
CGGAACACACGCATCTACAAGGACACCTCACAGAATTTCCAAGTGAGCCTGAATGTCTTGC GCGGCTACTACAACCAGAGCG
AGGCCGGTCTCACACCTACCAGTGGCTTTGTGGATGCTATGTTGCGCGGGACGGGCGCCTCCTCCGCGGGTACAGTCAGTT
CGCCTATGACGGCGCGGATTACATCGTCCTGAACGAGGACCTGCGCTCCTGGACCGCGTGGACATGGCGGCTCAGATCACT
CGGCGCAAATGGGAGGAGGAACTGTGGCAGAGCAGAGCAGAGCCTACCTGGAAGTGGCATGCGTGCAGTCGCTCCACAGAT
ACCTGGTGAACGGGAAGGAGACGCTACAGCGGTTCAGATCCTCCAAAGACACATGTGACCCGCCACCCAGCTCTGACATAA
AGTCACCTTGAGGTGCTGGGCCCTGGGCTTCTACCCTAAGGAGATCTCCCTGACCTGGCAGCAGGAGGGGAGGACCAGAGC
CAGGACGTGGAGTTGTGGAAACCAGGCCCTCAGGGGATGGGACCTTCCAGAAGTGGGCGGCCCTGGTGGTGCCTCCTGGAG
AGGAGCAGAGCTACACCTGCCATGTGCAGCAGGAGGGCCTGCAGGAGTCCCTCACCTGAGATGGGACCTCCTCAGCCCC
CGTCCCCATCGTGGGCATCATTGTTGGCCTGGTTCTCGTCTGCTGGTGGAGCCGTGGTGACTGGAGTTGTGATCTGGAGG
AAGAAGTGCTCAGGTGGAAAAGTAAGGAAATACCAACAGGCTGAAGGAGCAGCCGTAGCGAGAATCTGGTGTATGTTTGA
TGCCTTTCAAAGCTGAGACCTGGAGGCCCTGAGTGGGAAACAGCTGCCTTATACAAGACGGACTGATCAAGTAACAAACAT
CTCCCTCATGGCACTG

> MeLiM485_SLA-8 1068 bp

ATGGAGTCTCAGATGCTTCTTCTGGTGCTCTTGGGGGCCCTGACCGAGACCTGGGCGGCTCTCACTCCATGAGGTATTTCC
ACACGGTCGTGTCCCGGCCCGGCCACGGGGAGCCCGGTACCTTGAGGTGCGCTACGTGGACGACACGCAGTTTCGTGCGGTT
CGACAGCGAAGCTCAGAATCCGAGGATGGAGCCGCGGGCGCGTGGGTGGAGCAGGAGGGGAGGAGTATTGGGATGAGGAG
ACACAGCGCGCCAAGGATCTTGTAACAGATTTTTCGAAGGAACCTGATGATCCTTCGGGGCTACTACAACCAGAGTGAGACCG

GGTCTCATACCTTCCAGTTGACCTATGGTTGCGAGGAGGGGTCGCACGGACGCCCCCTTCACGCGCACTGGCAGTACGCCTA
CGAGGGCGAGGATTACATCACTTTGAACGAGGACCTGAGTTCCTGGACTGCGGCGGACATGGCGGCTCGGGTCACACAGCGC
AAGTGGGACAAGTCCCGTGCTCATGAGCGCTTCAAGAGCTACCTGGAGGGCACGTGTGTGGAGTGGCTCAGAAGATACCTGG
AGAACGGGAGGGAGATGCTGCAACGTGCAGACCCCCAAAGGCATATGTGACCTGCCACCCCAGCTCTGACAATAAAGTCAC
CCTGAGGTGCTGGGCCCTGGGCTTCTACCCCTAAGGAGATCTCCCTGACCTGGCGGCGGGAGGGGCAGGACCAGAGCCAGGAC
GTGGAGGTGTGTGAAACCAGGCCCTCAGGGGATGGGACCTTCCAGAAAGTGGGCGGCCCTGGTGGTGCCTCCTGGAGAGGAGC
AGAGCTACACCTGCCATGTGAAGCACGAGGGCCTGCAGGAGCCACTCACCTGAGATGGGAACCGTCTCGGCTGTCCGCCAT
CACCATCGTAGGAATCGTTGCTGGTCTGGTTCCTTGAGCTGTGATCACTGTGATATGGAAAAAAGGTTCTCAGGCAGA
AAAAGAGGAAGCTACTCTCAGGCGCCAAACAACAACAGCGTCGAGAACTCTGATGTGTCTCCCGAAAGTCCTCAAGGCATCT
GA

> MeLiM485_B2M 357 bp

ATGGCTCCCCCTCGTGGCCTTGGTCCCTGCTCGGGCTGCTCTCACTGTCTGGCCTGGATGCGGTGCGCGTCCCCGAAGGTTT
AGGTTTACTCACGCCACCCAGCGGAAAACGGAAAGCCAAATTACCTGAACTGCTATGTATCTGGGTTCATCCGCCCCAGAT
TGAAATTGATTTGCTGAAAAACGGGGAGAAGATGAACGCGGAGCAGTCAGACCTGTCTTTCAGCAAGGACTGGTCTTTCTAC
CTTCTGGTCCACACTGAGTTCACTCCTAACGCTGTGGATCAGTATAGCTGCCGCGTGAAGCACGTGACTCTCGATAAGCCCA
AGATAGTTAAGTGGGATCGAGACCACATAA

Supplementary file S2: Translated sequences of the cDNAs cloned into expression vectors with exon delineation for the SLA proteins

```
> MeLiM_SLA-1 361 AA 8 exons
MGPGALLFLLLSGTLLALTGTQAGPHSLSYFYTAVSRPDRGDSRFIAVGYVDDTQFVRFDSDAPNPRMEPRAPWIQQEGQDYWD
RETRKQRDTSQTYRVGLKNLRGYYNQSEAGSHTYQSMYGCYLGPDGLLLRGYRQYAYDGADYIALNEDLRSWTAADTAAQIT
KRKWETANVAERRRSYLQGLCVESLREYLEMGKDTLQRAEPPKTHVTRHPSSDLGVTLRCWALGFYPKEISLTGQREGQDQS
QDMELVETRPSGDGTFQKWAALVVPPGEEQSYTCHVQHEGLQEPLTLRWDPPQPPPVIVGIIVGLVLVLVAGAMVAGVVIWR
KKRSGEKGGSYTQAAGSDSDQGSDVSLTKGPRV

> MeLiM485_SLA-6 369 AA 7 exons
MQVTEPRTLLLVLSGSLALTETWAGHLHHGPKDPRLEKGQTRPQPLPVIGSHSLRYLHLIVSRPGHGSDYSSVGFLDDTQF
VRFSSDAANPRVEPRAPWMEQEGREYWDRQTDIAEHSKASRSNLRVIIGNHNHSQSESHSFLWVSGCDVGSDGRILRGYEQ
FSYDGDDYIVLNEDLRSWTAISTVAQIIRRKWEAEGVAEQYRAYLEIECVEWLRKYLEKGKDVLQRAVPPKTHVTRHPFYDN
KVTLRCWALGFYPKEISLTWQRDGEDQTQDMELVETRPSGDGTFQKWAALVVPSGEEQSYTCQVQHEGLQEPLLRWEPPPL
SLNLIMICIPVSVLVVTVLGTVIWRKRNSGGNRGNYVQAAS

> MeLiM485_SLA-7 388 AA 8 exons
MGPQALLLLLLSGTLVLTQPWARPHSLRYFYTAVSRPSRRDPRFSVVGYVDDTQFVRFDSNAPNPREEPRTPWVELGPEYCD
RNTRIYKDTSQNFQVSLNVLRGYYNQSEAGSHTYQWLCGCYVARDGRLLLRGYSQFAYDGADYIVLNEDLRSWTAVDMAAQIT
RRKWEEETVAEQSRAYLEVACVQSLHRYLVNGKETLQRSDPPKTHVTRHPSSDNKVTLRCWALGFYPKEISLTWQQEGQDQS
QDVEVVETRPSGDGTFQKWAALVVPPGEEQSYTCHVQQEGLQESLTLRWDPPQPPVIVGIIVGLVLVLVAGAVTGVVIWR
KKCSGGKVRKYQQAEGSSRSENSGVCLMPFKAETLEALSGKQLPYTRRTDQVTNISLMAL

> MeLiM485_SLA-8 355 AA 8 exons
MESQMLLLVLLGALTETWAGSHSMRYFHTVVSRPGHGEPRYLEVGYVDDTQFVRFDSEAQNPRMEPRAPWVEQEGQEYWDEE
TQRAKDLVQNFRRNLMILRGYYNQSETGSHTFQLTYGCEEGSHGRPLHAHWQYAYEGEDYITLNEDLSSWTAADMAARVTQR
KWDKSRAHERFKSYLEGTCVEWLRRYLENGREMLQRADPPKAYVTCHPSSDNKVTLRCWALGFYPKEISLTWREGQDQSQD
VEVETRPSGDGTFQKWAALVVPPGEEQSYTCHVKHEGLQEPLTLRWEPSRLSAITIVGIVAGLVLLGAVITVIWKKRFSGR
KRGSYSQAPSNNSVENSDVSPESPQGI

> MeLiM485_B2M 118 AA
MAPLVALVLLGLLSLSGLDAVARPPKVQVYSRHPAENGKPNYLNCYVSGFHPPQIEIDLLLKNGEKMNAEQSDLSFSKDWSFY
LLVHTEFTPNAVDQYSCRVKHVTLDKPKIVKWDRDH
```



STUDY REPORT

GENETIC IMMUNIZATION AGAINST-SWINE HMC PROTEINS:SLA6, SLA7 AND SLA8

TEST FACILITY :

IN-CELL-ART

1 place Alexis Ricordeau

44093 Nantes, France

SPONSOR

Laetitia MENDES

Ingénieur Commercial

Société P.A.R.I.S

21 Quai du Clos des Roses

60200 Compiègne, France

IN CELL ART SCIENTIFIC RESPONSIBLES

Benoît BARTEAU

Bruno PITARD

Study reference: ICA09-133

Originals of approved report:

Copies of approved report:

Study Timetable:

Strat of experimental phase: 18 August 2009

End of experiment phase: on-going

First draft study report: 16 December 2009

Report final version:

OBJECTIVES

The objective of this study was to produce, by DNA vaccination, mouse monoclonal antibodies against 3 swine CMH proteins named SLA6, SLA7 and SLA8. Immunizations were performed by IN CELL ART with SLA6, SLA7 or SLA8 encoding plasmid using *ICAntibodies*® technology. Humoral immune response was monitored by Western Blot. Cellular fusion and hybridomas production will be performed by P.A.R.I.S.

Study PROTOCOL

Animal model

20 Balb/C female mice (JANVIER, Le Genest Saint Isle, France), 8 weeks old were *i.m.* injected with *ICAntibodies*® formulation technology.

Animal housing

Animals were housed in plastic boxes of standard dimensions for housing mice. Animals were placed in an air-conditioned (15 - 21°C) environment. The artificial day/night light cycle involved 12 hours light and 12 hours darkness with light on at 8:00 a.m.

Vaccination protocol and study design

Study design

The 20 mice were divided in 4 groups:

Group #1: mice (n°1 to 8) were poly-immunized, *i.e. i.m.* injected with the all three SL6, SLA7 and SLA8 encoding plasmids associated with *ICAntibodies*® formulation technology.

Group #2: 4 mice (n°9 to 12) were *i.m.* injected with SL6 encoding plasmid associated with *ICAntibodies*® formulation technology.

Group #3: 4 mice (n°13 to 16) were *i.m.* injected with SL7 encoding plasmid associated with *ICAntibodies*® formulation technology.

Grpupe #4: 4 mice (n°17 to 20) were *i.m.* injected with SL8 encoding plasmid associated with *ICAntibodies*® formulation technology.

Of Note: mice n°10, 11 and 16 died during immunization experiment

Vaccination scheme

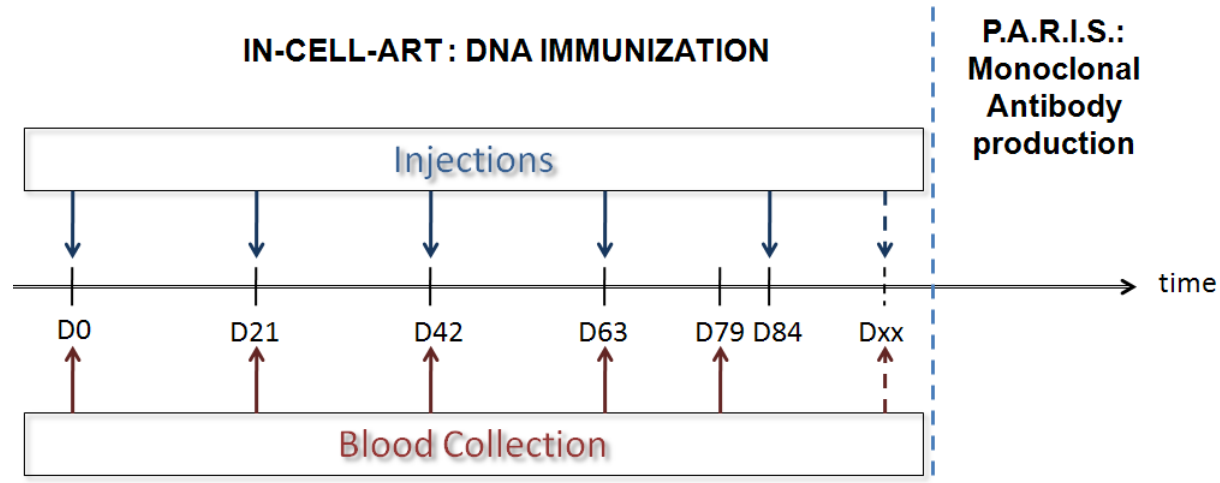


Figure 1: Project scheme

Blood samples were collected on day 0, 21, 42, 63 and 79. Each animal was i.m. injected on day 0, 21, 42, 63 and 84. Few days before splenocytes fusion for monoclonal antibody production, an additional injection will be performed.

Material AND methods

Plasmid

The three SLA sequences were synthesized and subcloned in pVAX plasmids containing kanamycin resistance gene.

The plasmid pVAX1-SLA6 (4050 bp) contains a cDNA insert of 1119bp encoding SLA6 gene (cloning site HindIII/XhoI) under the control of a CMV promotor (Figure2).

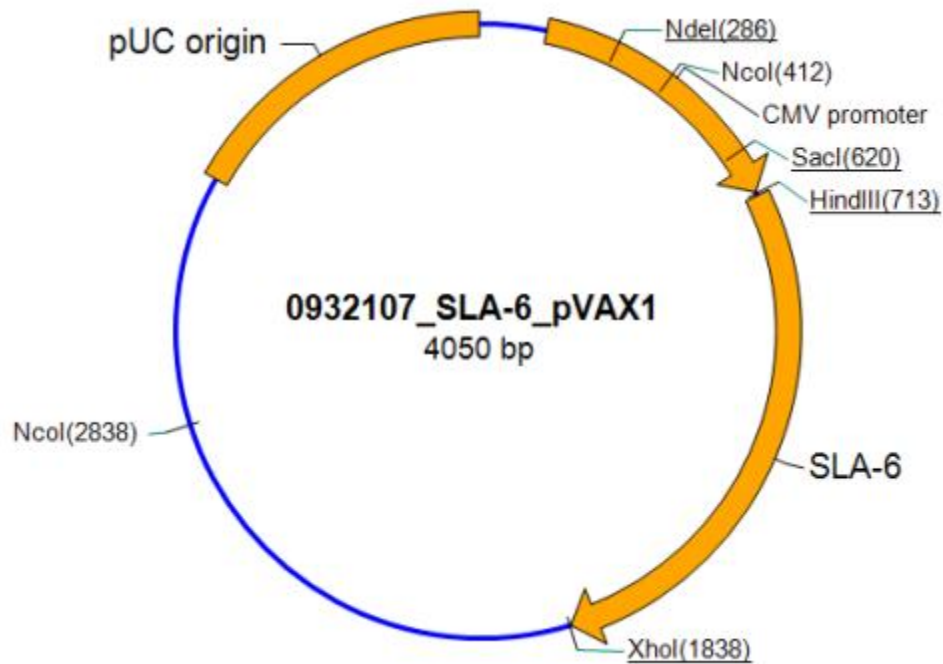


Figure 2: Restriction map of plasmid pVAX1-SLA6

The plasmid pVAX1-SLA7 (4107 bp) contains a cDNA insert of 1176bp encoding SLA7 gene (cloning site HindIII/XhoI) under the control of a CMV promoter (Figure 3).

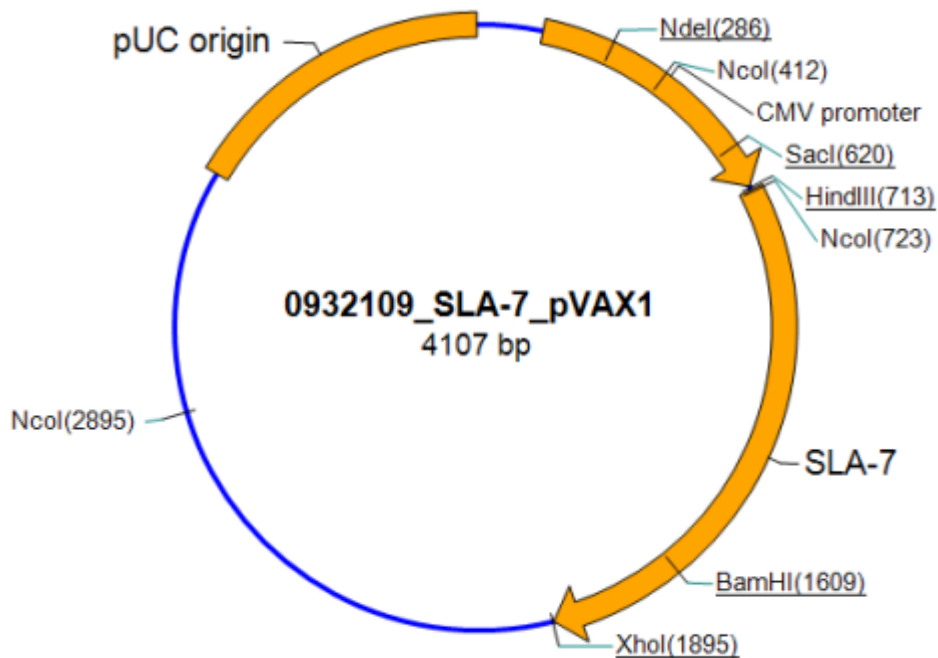


Figure 3: Restriction map of plasmid pVAX1-SLA7

The plasmid pVAX1-SLA8 (4008 bp) contains a cDNA insert of 1077bp encoding SLA8 gene (cloning site HindIII/XhoI) under the control of a CMV promoter (Figure 4).

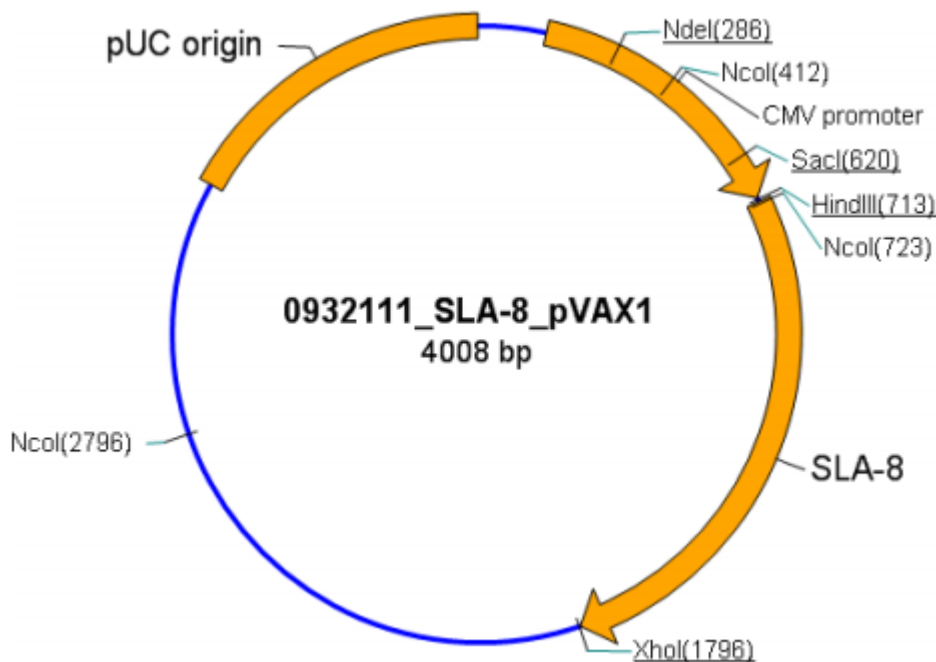


Figure 4: Restriction map of plasmid pVAX1-SLA8

SLA gene and corresponding protein

The encoded SLA genes belong to Sus Scrofa

Gene name

Sus scrofa SLA-6

Sus scrofa SLA-7

Sus scrofa SLA-8

Origin

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.

Nucleotide sequence

SLA 6

ATGCAGGTCACGGAGCCTCGAACCCTCCTGCTGGTGCTCTCGGGGTCCCTGGCCCTGACCGA
GACCTGGGCAGGCCACCTTCACCACGGACCCAAGGACCCGCGTCTGGAGAAGGGTCAGACCC
GGCCTCAGCCCCCTCCCTGTTATAGGATCCCACTCGCTAAGATACCTCCATATTTTGGTATCC

CGGCCCCGGCCACGGTAGTGACCTTTACAGCTCTGTGCGTTTCTTGACGACACGCAGTTCGT
ACGGTTCAGCAGCGATGCCGCGAATCCAAGGGTGGAGCCTCGGGCGCCGTGGATGGAGCAGG
AGGGGCGGGAATATTGGGATCGTCAGACTGATATAGCCAAAGAACATTCTGAAGGCTTCAAGA
TCGAACCTGCGGGTGATCATTGGCAACCACAACCATAGCCAGTCGGAGTCGCACAGCTTTCT
TTGGGTATCCGGCTGCGACGTGGGATCCGACGGGCGCATCCTTCGAGGGTACGAGCAGTTCT
CCTACGACGGCGACGATTACATCGTCCTGAATGAAGACCTGCGCTCCTGGACCGCGATCAGT
ACAGTGGCTCAGATCATCCGGCGCAAGTGGGAGGCGGAGGGAGTAGCTGAACAGTATCGGGC
ATACTTGAGATCGAGTGTGTGGAGTGGCTCCGCAAGTACCTCGAGAAGGGGAAGGACGTGC
TGCAGCGCGCAGTTCCCTCCAAAGACTCATGTGACCCGCCACCCCTTCTATGACAATAAAGTC
ACTCTAAGATGCTGGGCCCTGGGCTTCTACCCTAAGGAGATCTCCCTGACCTGGCAGCGTGA
TGGGGAGGACCAGACCCAGGACATGGAGCTCGTGGAGACCAGGCCTTCAGGGGATGGGACCT
TCCAGAAGTGGGCGGCCCTGGTGGTGCCTTCTGGAGAGGAGCAGAGCTACACCTGCCAGGTG
CAGCACGAGGGCCTGCAGGAGCCCCTCACCTGAGATGGGAGCCACCTCCTCTGTCCCTCAA
CCTCATCATGATTTGCATTCTGTGAGTGTGCTTGTGGTCACTGTGCTTGAACTGTGATCT
GGAGAAAGAGGAACTCAGGTGGAAATAGAGGGAACCTATGTTCAAGCTGCAAGCTGA

SLA7

ATGGGGCCCCAAGCCCTCCTCCTGCTGCTCTCCGGGACCCTGGTATTGACCCAGCCCTGGGG
ACGTCCCCACTCCCTGAGGTATTTCTACACAGCCGTGTCCCGACCGAGTCGACGGGATCCCC
GCTTCTCCGTCGTCGGCTATGTAGACGACACGCAGTTCGTGCGGTTTCGACAGCAATGCCCCG
AATCCGAGGGAGGAGCCGCGGACACCGTGGGTGGAGCTGGAGGGCCCGGAGTATTGCGATCG
GAACACACGCATCTACAAGGACACCTCACAGAATTTCCAAGTGAGCCTGAATGTCTGCGCG
GCTACTACAACCAGAGCGAGGCCGGGTCTCACACCTACCAGTGGCTTTGTGGATGCTATGTT
GCGCGGGACGGGCGCCTCCTCCGCGGGTACAGTCAGTTTCGCCTATGACGGCGCGGATTACAT
CGTCCTGAACGAGGACCTGCGCTCCTGGACCGCGGTGGACATGGCGGCTCAGATCACTCGGC
GCAAATGGGAGGAGGAAACTGTGGCAGAGCAGAGCAGAGCCTACCTGGAAGTGGCATGCGTG
CAGTCGCTCCACAGATACCTGGTGAACGGGAAGGAGACGCTACAGCGGTCAGATCCTCCAAA
GACACATGTGACCCGCCACCCAGCTCTGACAATAAAGTCACCTTGAGGTGCTGGGCCCTGG
GCTTCTACCCTAAGGAGATCTCCCTGACCTGGCAGCAGGAGGGGCAGGACCAGAGCCAGGAC
GTGGAGGTTGTGGAAACCAGGCCCTCAGGGGATGGGACCTTCCAGAAGTGGGCGGCCCTGGT
GGTGCCTCCTGGAGAGGAGCAGAGCTACACCTGCCATGTGCAGCAGGAGGGCCTGCAGGAGT
CCCTCACCTGAGATGGGACCCTCCTCAGCCCCCGTCCCCATCGTGGGCATCATTGTTGGC
CTGGTTCTCGTCCTGGTCGCTGGAGCCGTGGTGACTGGAGTTGTGATCTGGAGGAAGAAGTG
CTCAGGTGGAAGAAGTAAGGAAATACCAACAGGCTGAAGGGAGCAGCCGTAGCGAGAAGTCTG
GTGTATGTTTGATGCCTTTCAAAGCTGAGACCCTGGAGGCCCTGAGTGGGAAACAGCTGCCT
TATACAAGACGGACTGATCAAGTAACAAACATCTCCCTCATGGCACTGTGA

SLA8

ATGGAGTCTCAGATGCTTCTTCTGGTGCTCTTGGGGGCCCTGACCGAGACCTGGGCGGGCTC
TCACTCCATGAGGTATTTCCACACGGTCGTGTCCCGGCCCGGCCACGGGGAGCCCCGGTACC
TTGAGGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTTCGACAGCGAAGCTCAGAATCCG
AGGATGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGGGGCAGGAGTATTGGGATGAGGAGAC
ACAGCGCGCCAAGGATCTTGTACAGAATTTTTCGAAGGAACCTGATGATCCTTCGGGGCTACT
ACAACCAGAGTGAGACCGGGTCTCATACCTTCCAGTTGACCTATGGTTGCGAGGAGGGGTGCG
CACGGACGCCCCCTTACGCGCACTGGCAGTACGCCTACGAGGGCGAGGATTACATCACTTT
GAACGAGGACCTGAGTTCCTGGACTGCGGCGGACATGGCGGCTCGGGTCACACAGCGCAAGT
GGGACAAGTCCCCTGCTCATGAGCGCTTCAAGAGCTACCTGGAGGGCACGTGTGTGGAGTGG
CTCAGAAGATACCTGGAGAACGGGAGGGAGATGCTGCAACGTGCAGACCCCCCAAAGGCATA
TGTGACCTGCCACCCCAGCTCTGACAATAAAGTCACCCTGAGGTGCTGGGCCCTGGGCTTCT
ACCCTAAGGAGATCTCCCTGACCTGGCGGCGGGAGGGGCAGGACCAGAGCCAGGACGTGGAG

GTTGTGGAAACCAGGCCCTCAGGGGATGGGACCTTCCAGAAGTGGGCGGCCCTGGTGGTGCC
TCCTGGAGAGGAGCAGAGCTACACCTGCCATGTGAAGCACGAGGGCCTGCAGGAGCCACTCA
CCCTGAGATGGGAACCGTCTCGGCTGTCCGCCATCACCATCGTAGGAATCGTTGCTGGTCTG
GTTCTCCTTGGAGCTGTGATCACTGTGATATGGAAAAAAGGTTCTCAGGCAGAAAAAGAGG
AAGCTACTCTCAGGCGCCAAGCAACAACAGCGTCGAGAACTCTGATGTGTCTCCCGAAAGTC
CTCAAGGCATCTGA

For quality assurance document, see ICA09-133_PARIS_SLA678_RapportSequencage document.

Peptide sequence

SLA6

MQVTEPRTL LLLVLSGSLALTETWAGHLHHGPKDPRLEKGQTRPQPLPVIGSHSLRYLHILVS
RPGHGS DLYSSVGFLDDTQFVRFSSDAANPRVEPRAPWMEQEGREYWDRQTDIAKEHSKASR
SNLRVIIGNHNHSQSESHSFLWVSGCDVGS DGRILRGYEQFSYDGDDYIVLNEDLRSWTAIS
TVAQII RRKWEAEGVAEQYRAYLEIECVWLRLKYLEKGKDV LQRAVPPKTHVTRHPFYDNKV
TLRCWALGFYPKEISLTWQ RDGEDQTQDMELVETRPSGDGTFQKWAALVVP SGEEQSYTCQV
QHEGLQEPLTLRWEPPPLSLNLIMICIPVSVLVVTVLGTVIWRKRNSGGNRGNYVQAAS

SLA7

MGPQALL LLLSGTLVLTQPWARPHSLRYFYTAVSRPSRRDPRFSVVG YVDDTQFVRFDSNAP
NPREEPRTPWVELEGPEYCDRNTRIYKDT SQNFQVSLNVLRGYYNQSEAGSHTYQWLCGCYV
ARDGRLLRGYSQFAYDGADYIVLNEDLRSWTA VDMAAQITRRKWEETVAEQSRAYLEVACV
QSLHRYLVNGKETLQ RSDPPKTHVTRHPSSDNKVT LRCWALGFYPKEISLTWQQEGQDQSQD
VEVVETRPSGDGTFQKWAALVVP PGEEQSYTCHVQQEGLQESLTLRWDPPQPPVPIVGII VG
LVLVLVAGAVVTGVVIWRKKCSGGKVRKYQQAEGSSRSENSGVCLMPFKAETLEALSGKQLP
YTRRTDQVTNISLMAL

SLA8

MESQMLLLVLLGALTETWAGSHSMRYFHTTVVSRPGHGEP RYLEVGYVDDTQFVRF DSEAQNP
RMEPRAPWVEQEGQEYWDEETQRAKDLVQNFRNL MILRGYYNQSETGSHTFQLTYGCEEGS
HGRPLHAHWQYAYEGEDYITLNEDLSSWTAADMAARVTQRKWDK SRAHERFKSYLEGTCVEW
LRRYLENGREMLQ RADPPKAYVTCHPSSDNKVT LRCWALGFYPKEISLTWRREGQDQSQDVE
VVETRPSGDGTFQKWAALVVP PGEEQSYTCHVKHEGLQEPLTLRWEPSRLSAITIVGIVAGL
VLLGAVITVIWKKRFSGRKRGSYSQAPSNN SVENS DVSPESPQGI

Protein size and mass

SLA6:

Size: 369 Amino Acids

Molecular weight: 42043 Dalton (from PROTEIN CALCULATOR v3.3)

SLA7:

Size: 388 Amino Acids

Molecular weight: 43942 Dalton (from PROTEIN CALCULATOR v3.3)

SLA8:

Size: 355 Amino Acids

Molecular weight: 40796 Dalton (from PROTEIN CALCULATOR v3.3)

Bacterial strain

Plasmids pVAX1-SLA were transformed by IN-CELL-ART in DH5 α TM competent bacteria, obtained from Invitrogen (Cergy Pontoise, France), according to manufacturer protocol.

Name and batch

DH5 α TM, Subcloning EfficiencyTM, Chemically Competent Cells, Batch n°394577 (Invitrogen, Cergy Pontoise, France).

Genotype

F- ϕ 80lacZ Δ M15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ .

Amplification and purification of pVAX1-SLA plasmids**Bacteria amplification**

Transformed bacteria were amplified by performing a 60 μ l preculture in 60 ml of LB medium, supplemented with 300 μ l kanamycin incubated overnight at 37°C. Bacterial preculture (20 ml) was then cultured for 24 hours in 2 l of LB medium supplemented with 6.25 ml kanamycin. The whole volume was then centrifuged at 4 °C (15 min, 5000 rpm).

pVAX1-SLA plasmids purification

Plasmid purification was performed using purification kit « EndoFree Plasmid Mega Kit » (Qiagen S.A., Courtaboeuf, France) according to manufacturer protocol.

Plasmid quality control**pVAX1-SLA plasmids concentration determination**

Plasmid concentration was evaluated by measuring absorbance (260 nm) of diluted plasmid solution (1/250) in triplicate. Results of the 3 purifications are presented in figure 5.

Report		Test Type	Nucleic Acid	Date/time	17/09/2009 13:35	Page #	1					
Sample ID	User ID	Date	Time	ng/ul	A260	A280	260/280	260/230	Constant	Cursor Pos.	Cursor abs.	340 raw
SLA6 1/250	Default	17/09/2009	13:31	39,89	0,798	0,486	1,64	2,21	50,00	230	0,362	0,006
SLA6 1/250	Default	17/09/2009	13:31	38,68	0,774	0,466	1,66	2,18	50,00	230	0,354	0,005
SLA6 1/250	Default	17/09/2009	13:32	37,49	0,750	0,452	1,66	2,17	50,00	230	0,346	0,013
SLA7 1/250	Default	17/09/2009	13:32	40,52	0,810	0,474	1,71	2,24	50,00	230	0,362	0,018
SLA7 1/250	Default	17/09/2009	13:33	42,14	0,843	0,500	1,68	2,25	50,00	230	0,375	0,014
SLA7 1/250	Default	17/09/2009	13:33	40,00	0,800	0,453	1,77	1,96	50,00	230	0,409	0,009
SLA8 1/250	Default	17/09/2009	13:34	45,20	0,904	0,511	1,77	2,23	50,00	230	0,405	0,011
SLA8 1/250	Default	17/09/2009	13:34	44,38	0,888	0,521	1,70	2,20	50,00	230	0,404	0,009
SLA8 1/250	Default	17/09/2009	13:35	43,86	0,877	0,523	1,68	2,21	50,00	230	0,397	0,011

Figure 5: Plasmid concentration determination by optical density measurements.

For each plasmid, mean concentration was calculated and plasmid solutions were stored at - 20°C.

pVAX1-SLA plasmid identity

Digestion

Plasmid DNA identity was determined by restriction enzymatic digestion. Plasmid was digested by restriction enzymes HindIII and XhoI. Expected fragment sizes are presented in Table 1.

Table 1: Expected fragments size following XhoI and HindIII digestion

Digestion	pVAX1-SLA6	pVAX1-SLA7	pVAX1-SLA8
Non digested (ND)	<4050 bp	<4107	<4008
Non digested (ND) + buffer 2	<4050 bp	<4107	<4008
XhoI	4050 bp	4107	4008
XhoI and HindIII	1119 and 2931	1176 and 2931	1077 and 2931

Agarose gel electrophoresis

DNA fragments were separated by migration on 1% agarose gel. Figure 6 shows that non digested pVAX1-SLA plasmid shows high content of supercoiled form.

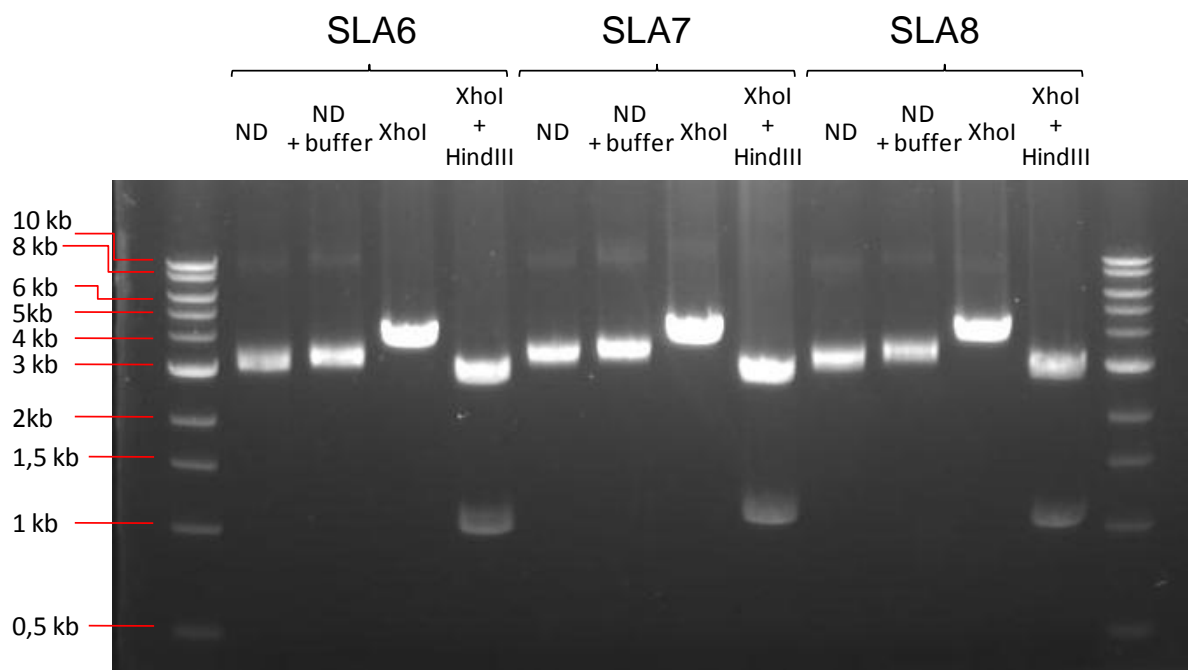


Figure 6: Plasmid pVAX1-SLA identity by agarose gel electrophoresis experiment.

Figure 6 also shows that plasmid digestion with XhoI resulted in the formation of one band of about 4000 bp. Plasmid digestion with the two XhoI and HindIII enzymes led to the formation of two bands of about 1000 and 3000 bp.

However, in order to better characterize the three pVAX1-SLA plasmids, more resolute gel electrophoresis experiment was performed as shown in figure 7.

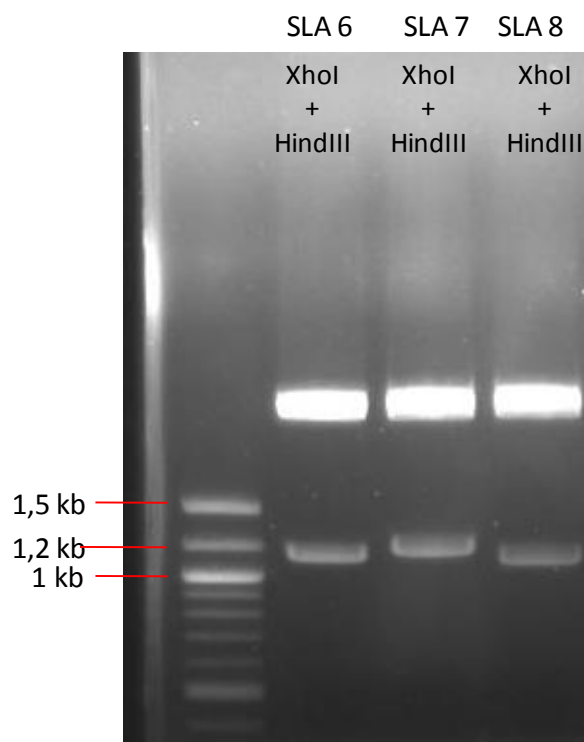


Figure 7: Plasmid pVAX1-SLA identity by agarose gel electrophoresis experiment

Gel electrophoresis of XhoI and HindIII digested pVAX1-SLA plasmids shows that the molecular weight of the small bands corresponded to the expected molecular weight, attesting that the purified plasmids are the expected ones.

Humoral immune response analysis

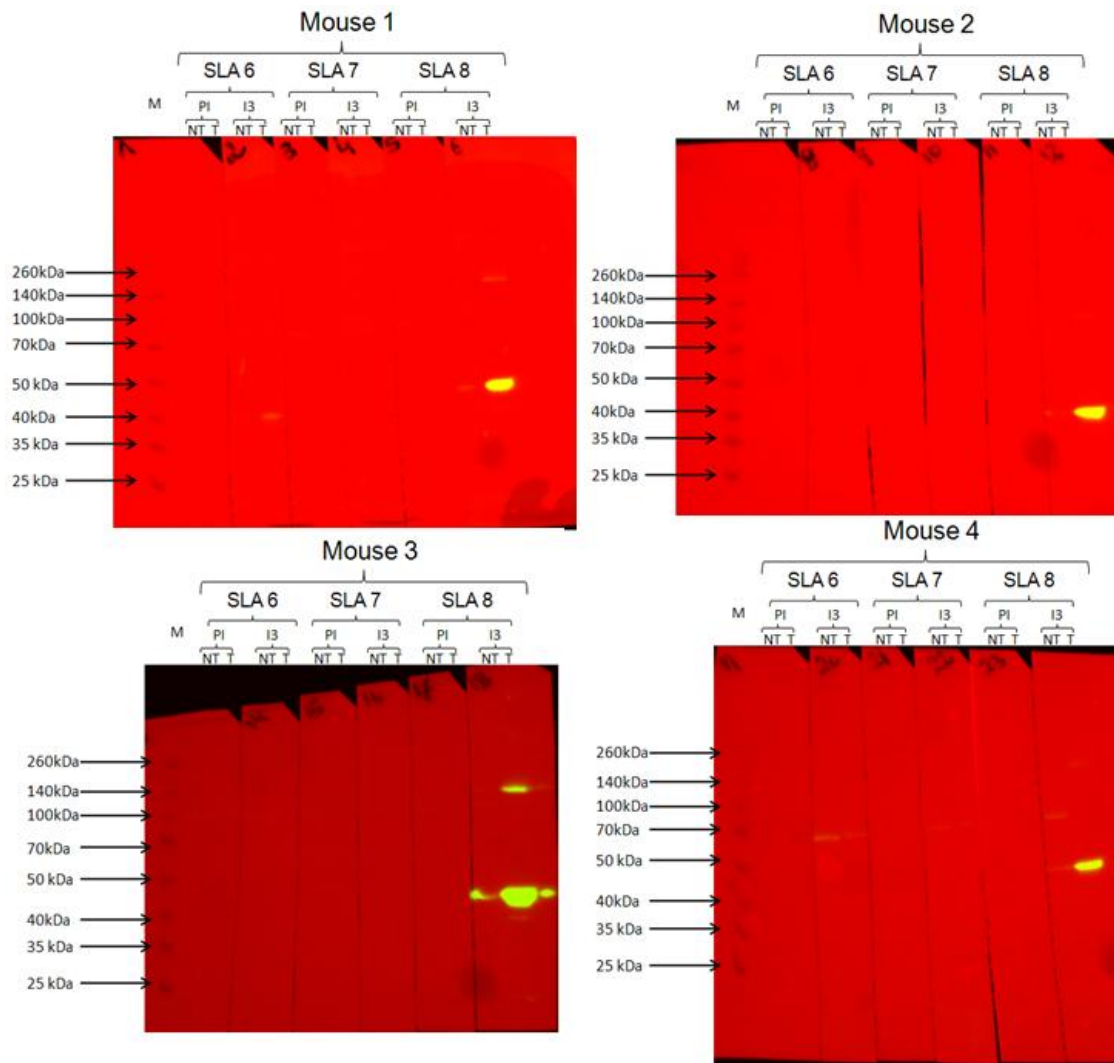
Western Blot

Presence of antibodies (total IgG) specific for SLA was analyzed by Western Blot. Briefly, lysate of pVAX1-SLA plasmid transfected or not HeLa cells with ICAfectin®441 (IN-CELL-ART, Nantes, France), was centrifugated. Then, supernatant was heat denatured at 95°C during 10 min. After SDS-PAGE migration of the supernatant, and nitrocellulose membrane transfer, 1/200 diluted serum were incubated with the membrane. Then, peroxidase-conjugated goat anti-mouse IgG, diluted at 1/5000, was incubated during one hour. After addition of the peroxidase substrate, luminescence was measured at different exposure times.

RESULTS

Anti SLA6-SLA7-SLA8 antibody production by poly genetic immunization

At day 63 after the primo-injection and 2 additional injections, western Blot analysis with ICAFectin®441 transfected HeLa cells shows that pre immune serum (PI) of the eight mice polyimmunized (SLA6-SLA7-SLA8) not react neither with the proteins encoded by the three pVAX1 plasmid nor with untransfected cell lysate (Figure 8)



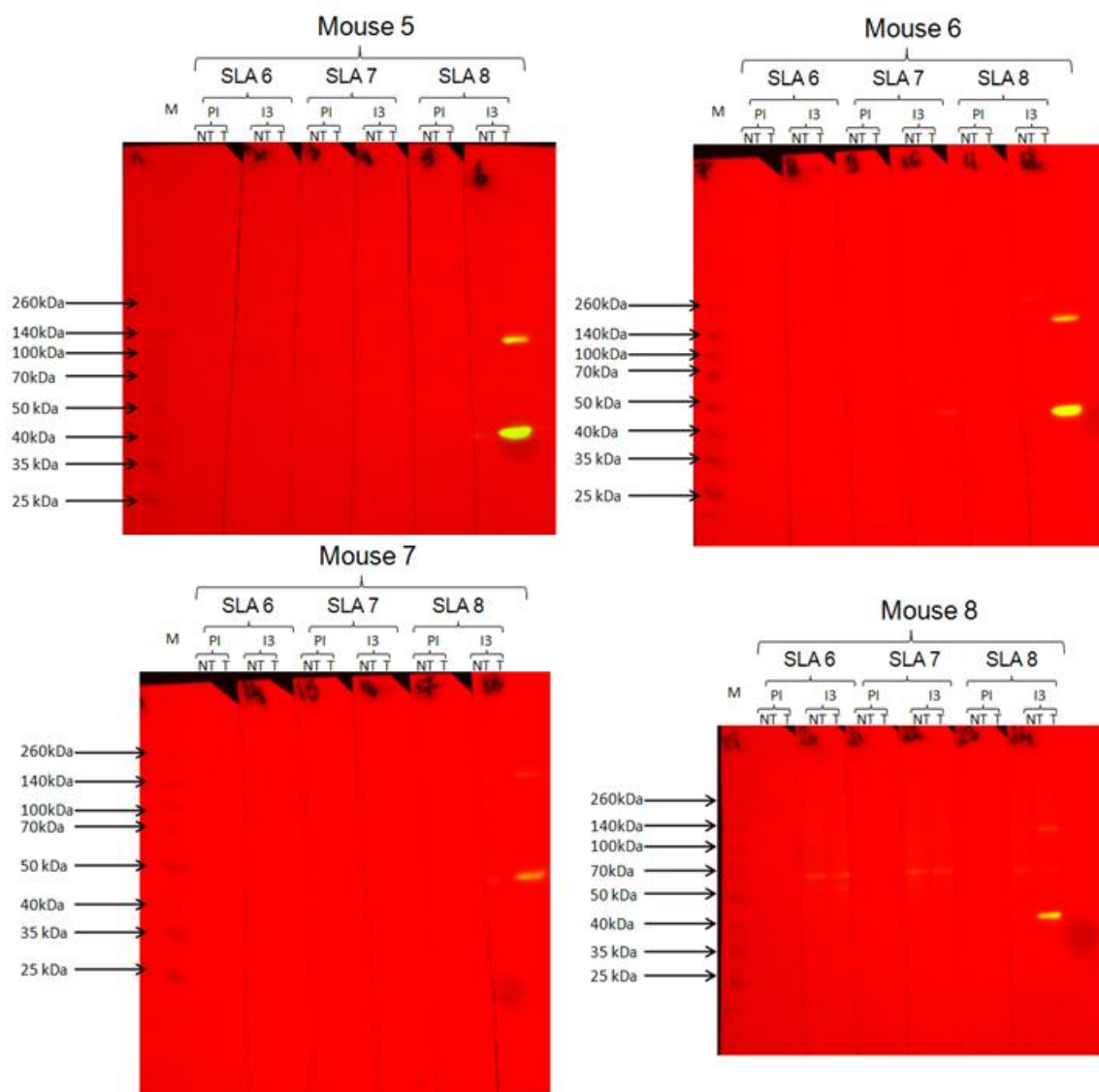


Figure 8: Western Blot analysis on lysate of pVAX1-SLA6 or 7 or 8/ICAFectin441 transfected (T) or not (NT) HeLa cells. For each poly-immunized mouse, pre-immune (PI) and D63 (I3) serum were tested.

At day 63 after the primo-injection and 2 additional injections, western Blot analysis shows that D63 serum of the eight poly-immunized mice reacted only with pVAX1-SLA8 transfected cell lysate. More precisely, all the poly-immunized mice serum recognized a protein which have a molecular weight around 40-45 kDa and only serum from mice n°3, 5 and 6 reacted with an additional 140 kDa molecular weight protein. The molecular weight of the band (40-45kDa) is compatible with SLA8 molecular weight (41 kDa). The 140 kDa molecular weight protein could correspond to a trimer of the 41 kDa SLA8 monomer.

Western blot analysis also shows that, probably due to immunogenicity differences between proteins, poly immunized mice only displayed anti SLA8 antibodies and these antibodies did

not cross react with SLA 6 and SLA7 proteins. Therefore, anti SL8 antibodies are highly specific of SL8 antigen.

Anti SL6-SLA7-SLA8 antibody production by mono genetic immunization

Figure 9 shows Western blot analysis of mono immunized mice either with pVAX1-SLA6 or pVAX1-SLA7 or pVAX1-SLA8 plasmid associated with *ICAntibodies* technology.

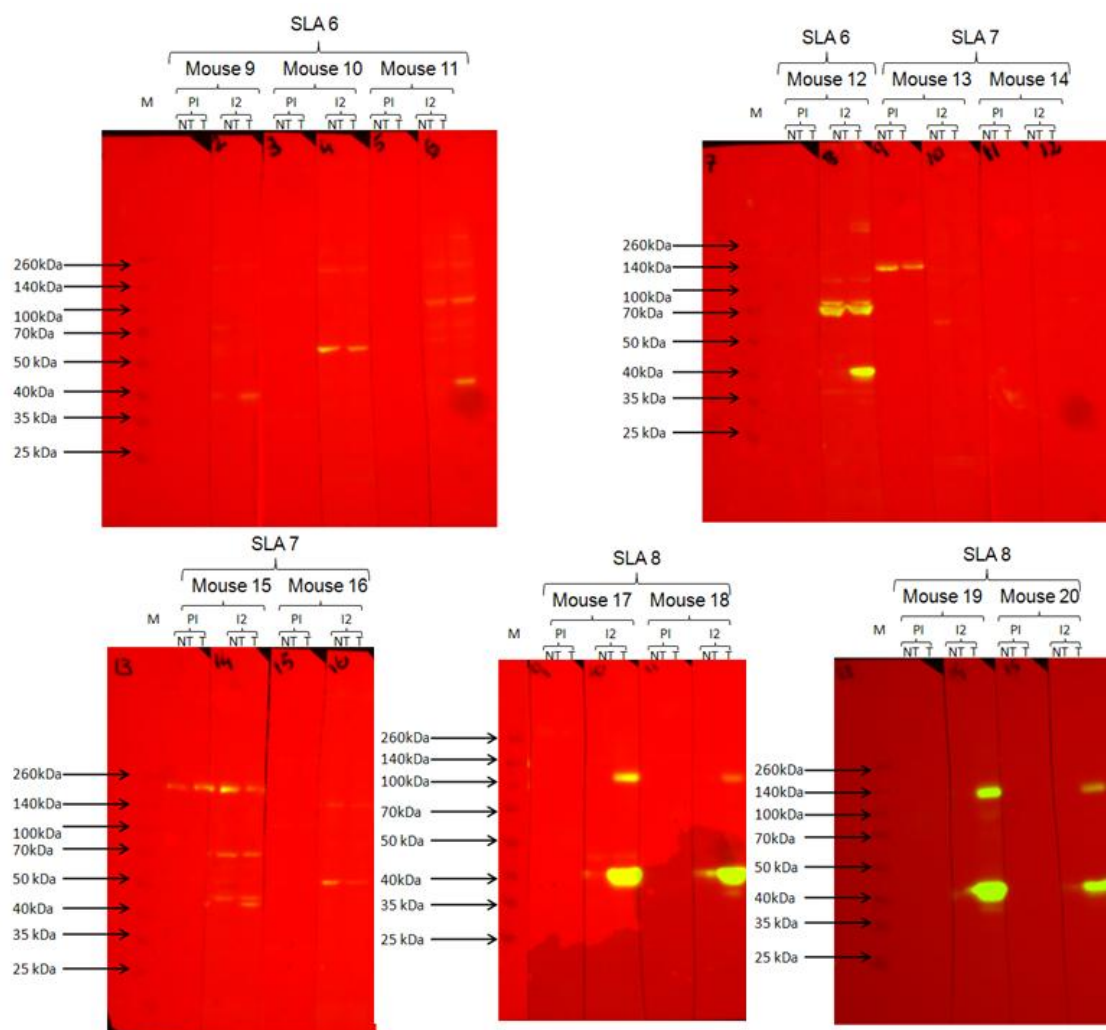


Figure 9: Western Blot analysis on lysate of pVAX1-SLA6 or 7 or 8/ICAFectin441 transfected (T) or not (NT) HeLa cells. For each mono-immunized mouse, pre-immune (PI) and D42 (I2) serum were tested.

At day 42 after primo injection and 1 additional injection, Western Blot analysis with ICAFectin®441 transfected HeLa cells shows that pre immune serum (PI) of the twelve mice, except mouse 13 and 15, did not react neither with the proteins encoded by the three pVAX1 plasmid nor with untransfected cell lysate (Figure 9).

Western Blot analysis shows that mice n°9, 11 and 12 immunized with 2 injections of pVAX1-SLA6 led to the production of antibodies targeted specifically against SLA6. Indeed, a 40kDa signal is only present with pVAX1-SLA6 transfected HeLa. Figure 9 also shows that serum from mouse 12 displayed a band which corresponds to a protein of 70 kDa molecular weight which is not specific of SLA6 as evidenced by the fact that this signal is also present with non transfected HeLa cells.

At day 42 after primo injection and 1 additional injection, western Blot analysis shows that SLA7 vaccinated mice were not or poorly immunized given that no specific signal can be observed.

At day 42 after primo injection and 1 additional injection, western Blot analysis shows that all mice immunized with plasmid encoding SLA8 has developed high titer of anti SLA8 antibodies as evidenced by the high 40kDa band intensity. As already observed with polyimmunized mice a 140kDa protein reacted with serum of mice monoimmunized.

Figure 10 shows Western blot analysis of mono immunized mice either with pVAX1-SLA6 or pVAX1-SLA7 plasmid associated with *ICAntibodies* technology, at day 79 after primo injection and 3 additional injections.

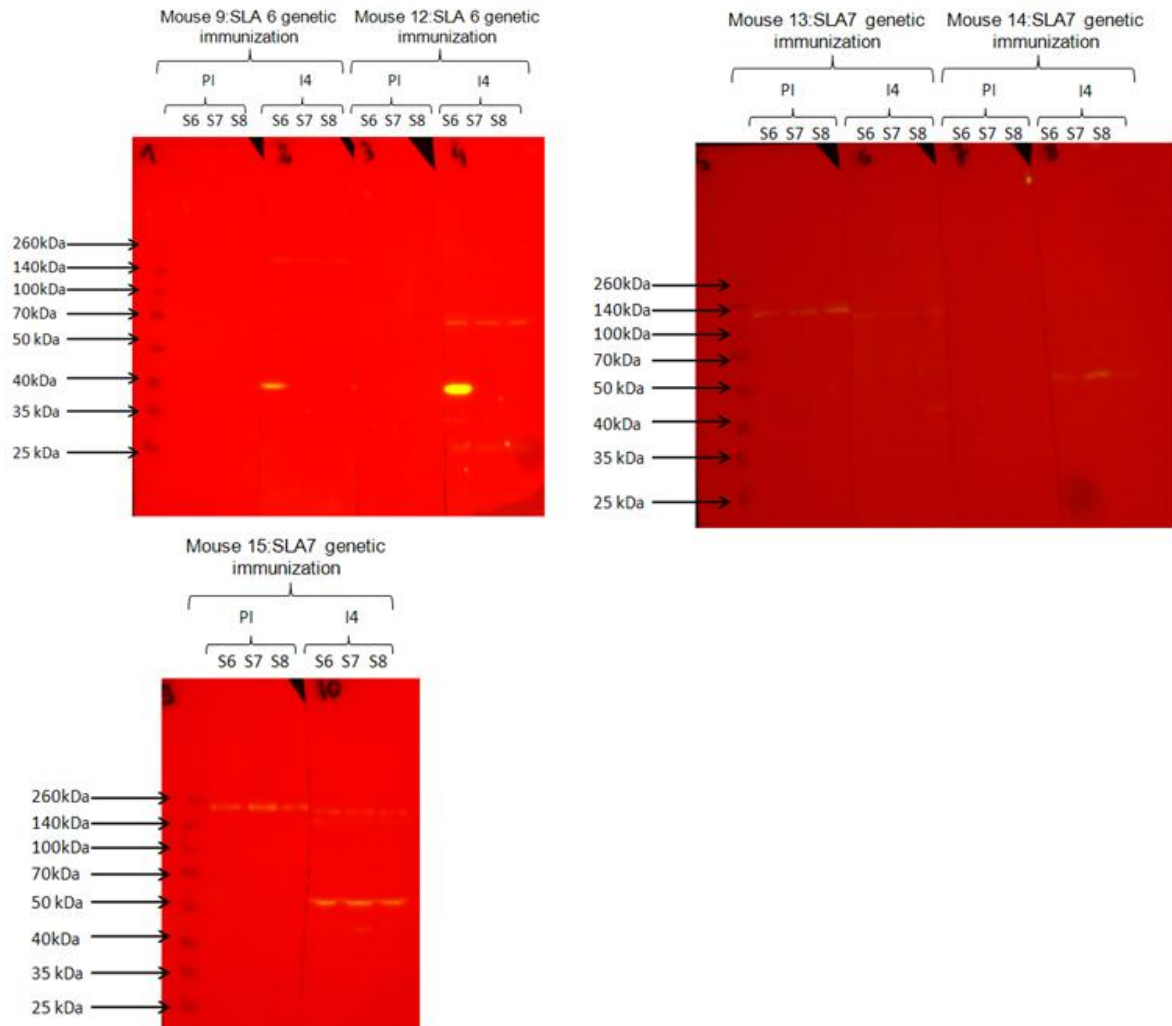


Figure 10: Western Blot analysis on lysate of pVAX1-SLA6 (S6) or 7 (S7) or 8 (S8)/ICAFectin441 transfected (T) HeLa cells. For each genetic mono-immunized mouse, pre-immune (PI) and D79 (I4) serum were tested.

Western Blot analysis shows that mice n°9 and 12 immunized with 4 injections of pVAX1-SLA6 led to the production of antibodies targeting specifically SLA6 and do not cross reacting neither with SLA7 nor SLA8.

At day79 after primo injection and 3 additional injections, western Blot analysis shows that SLA7 vaccinated mice were not or poorly immunized given that no specific signal can be observed.

Thus, results obtained at day 79, after 4 injections, confirmed those obtained at day 42 after 2 injections, for SLA6 and SLA7 genetic immunized mice.

Anti SLA6-SLA7-SLA8 antibody production by peptidic immunization

Next, we investigated whether serum (provided by the sponsor) of mice immunized with 2 peptides for each SL6, SLA7 and SLA8 antigens could recognize the whole SLA6, SLA7 or SLA8 protein by Western Blot analysis. To this end, western Blot experiments were performed on lysate of pVAX1-SLA6, pVAX1-SLA7 or pVAX1-SLA8 transfected HeLa cells with *ICAFectin*®441 (Fig.11).

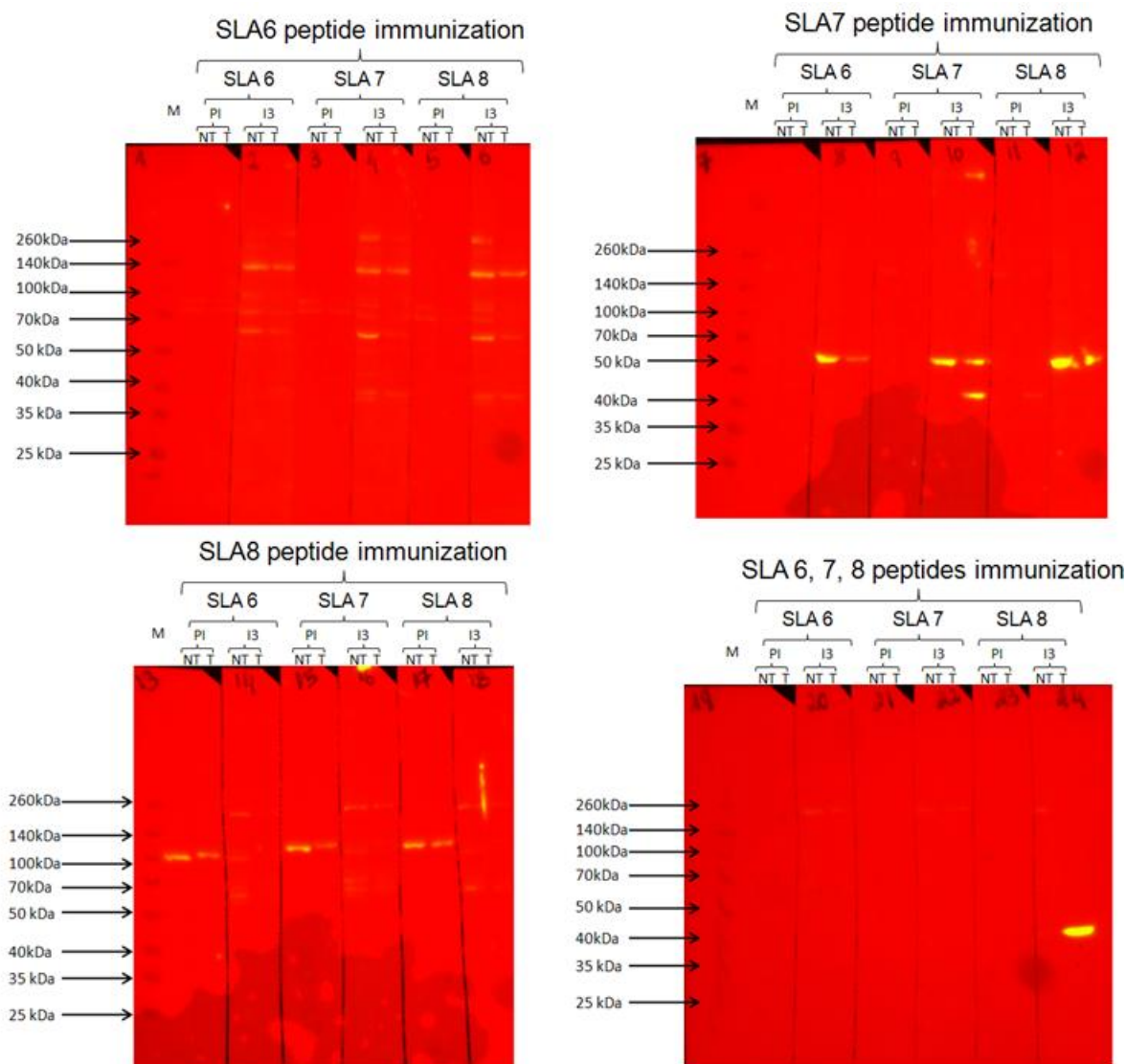


Figure 11: Western Blot analysis on lysate of pVAX1-SLA6 or 7 or 8/ICAFectin441 transfected (T) or not (NT) HeLa cells. For each mono or poly peptide-immunized mouse, naive (PI) and DXX (I3) serum were tested.

Western Blot analysis shows that serum of mice immunized three times with two peptides of SLA7, reacted specifically with pVAX1-SLA7 transfected HeLa cells lysate. The 40kDa molecular weight of the band is compatible with the 44 kDa molecular weight of the SLA7 protein.

By contrast, western blot analysis of serum of mice immunized with three injections of 2 peptides corresponding of SLA6 and SLA8 antigens did not show specific antibodies targeted against SLA6 and SLA8.

Surprisingly, western Blot analysis of serum of mice immunized with three injections of 2 peptides of each SLA6, SLA7 and SLA8 antigens (6 peptides in total) shows that the poly peptide-vaccinated mice were strongly immunized against SLA8 protein as evidence by the high 40kDa band intensity, while no immunization was detected for mono SLA8 peptide vaccinated mice.

CONCLUSION and future directions

Here we showed by western blot analysis that mouse mono-vaccination with non viral vectors encoding swine whole CMH proteins named SLA6, SLA7 and SLA8 allowed generation of anti SLA 6 and SLA 8 antibodies. Western blot analysis of serum of mouse immunized with SLA7 did not allow to show antibodies against SLA7, however we can not exclude that conformational antibodies against SLA7 were produced but could not been seen by WB due to the denatured form of SLA7 on the gel.

Of note, poly immunized mice with the three plasmids only displays anti SLA8 antibodies and these antibodies do not cross react with SLA 6 and SLA7 proteins. This could be related to immunogenicity differences between proteins

Peptide immunized mice serum generated by the sponsor recognize, in Western Blot experiment, SLA 7 and SLA8 proteins for mono and poly peptide vaccinated mice respectively.

ABSTRACT

Analysis of the non classical class I genes of the MHC in swine

In pig, very little information is available on the three non classical MHC class I genes SLA-6, -7 and -8 (SLA-Ib genes). Our aim was to increase knowledge on SLA-Ib genes by studying their polymorphism, transcription and protein expression. Full length transcripts were characterized from thymus and brain of MeLiM pigs resulting in the annotation of 8 exons for SLA-7 and -8 and 7 exons for SLA-6. The three full length cDNAs encode molecules with a predicted folding consistent with peptide presentation. No additional transcript was found for SLA-8 while four and five were detected for SLA-6 and SLA-7, respectively, suggesting that some soluble forms of these molecules may exist. For SLA-7, an alternative spliced variant was found in the 3'UTR of the gene after the termination codon suggesting possible post-transcriptional regulation of the gene. Polymorphism studies confirmed a limited nucleotide polymorphism but revealed an unexpected copy number variation for SLA-7. Ongoing experiments are still in progress to characterize monoclonal antibodies specific of each SLA-Ib molecule in transient transfection systems. Our overall results provided significant new data on SLA-Ib genes and will pave a way toward more functional analyses related to their putative function in immuno-tolerance.

RESUME

Etude des gènes de classe I non classiques du complexe majeur d'histocompatibilité chez le porc

Chez le porc, les trois gènes du CMH de classe I non classiques (gènes SLA-Ib) SLA-6, -7 et -8 sont peu étudiés alors que ce sont des candidats majeurs comme homologues fonctionnels des gènes humains HLA-E, -F et -G qui font l'objet de travaux soutenus de par leur rôle dans la tolérance fœto-maternelle lors de la grossesse chez la femme notamment. Notre objectif a consisté à analyser le polymorphisme, la transcription et l'expression des protéines de ces trois gènes d'intérêt. Nous avons montré que les trois gènes sont transcrits dans une large gamme de tissus, avec une expression prépondérante dans les tissus lymphoïdes, le système digestif et les poumons. Les trois gènes expriment des transcrits qui codent pour des protéines de pleine longueur dont la traduction et la prédiction de conformation sont compatibles avec la présentation de peptides à la surface cellulaire. Les gènes SLA-6 et -7 expriment des transcrits alternatifs qui pourraient coder pour des isoformes solubles, alors qu'un unique transcrit a été trouvé pour le gène SLA-8. Un épissage alternatif dans la région 3' non codante en aval du codon de terminaison a également été identifié pour SLA-7, suggérant l'existence de mécanismes de régulation post-transcriptionnelle pour ce gène. Les analyses de polymorphisme ont confirmé des variations nucléotidiques limitées mais mis en évidence une variation du nombre de copies du gène SLA-7 selon les animaux. Des analyses sont encore en cours pour caractériser la spécificité d'anticorps monoclonaux dirigés contre les molécules SLA-Ib dans des expériences de transfections transitoires en cultures de cellules. Nos résultats sont une contribution importante à la caractérisation des gènes SLA-Ib chez le porc et permettront de poursuivre avec des approches plus fonctionnelles visant l'analyse de leurs possibles fonctions dans des mécanismes liés à l'immuno-tolérance.